USP7 small-molecule inhibitors interfere with ubiquitin binding

Kategaya, Lorna; Di Lello, Paola; Rougé, Lionel; Pastor, Richard; Clark, Kevin R.; Drummond, Jason

Published in:
Nature

DOI:
10.1038/nature24006

Publication date:
2017

Document Version
Peer reviewed version

USP7 Small Molecule Inhibitors Interfere with Ubiquitin Binding

Lorna Kategaya[1,2†], Paola Di Lello[3†], Lionel Rougé[3], Richard Pastor[4], Kevin R. Clark[5], Jason Drummond[5], Tracy Kleinheinz[5], Eva Lin[1], John-Paul Upton[1,2], Sumit Prakash[1,2], Johanna Heideker[1,2], Mark McCleland[1,2], Maria Stella Ritorto[6], Travis W. Bainbridge[7], Taylur P. Ma[8], Michael C.M. Kwok[7], Zachary Stiffler[9], Bradley Brasher[9], Dario R. Alessi[6], Matthias Trost[6], Adam R. Renslo[10], Michelle R. Arkin[10], Frederick Cohen[4], Kebing Yu[8], Frank Peale[11], Florian Gnad[12], Matthew T. Chang[12], Christiaan Klijn[12], Elizabeth Blackwood[13], Scott E. Martin[1], William F. Forrest[12], James A. Ernst[7], Chudi Ndubaku[4], Xiaojing Wang[4], Maureen H. Beresini[5], Vickie Tsui[4], Carsten Schwerdtfeger[9], Robert A. Blake[5], Jeremy Murray[3], Till Maurer[3]* and Ingrid E. Wertz[1,2]*

Departments of [1]Discovery Oncology, [2]Early Discovery Biochemistry, [3]Structural Biology, [4]Discovery Chemistry, [5]Biochemical and Cellular Pharmacology, [7]Protein Chemistry, [8]Microchemistry, Proteomics, and Lipidomics, [11]Research Pathology, [12]Bioinformatics and Computational Biology, and [13]Translational Oncology, Genentech, South San Francisco, CA 94080, USA.
[6] MRC Protein Phosphorylation and Ubiquitylation Unit, School of Life Sciences, University of Dundee, Dundee DD1 5EH, UK.
[9] Boston Biochem, Inc. 840 Memorial Drive, Cambridge, MA, USA.
[10] Department of Pharmaceutical Chemistry and Small Molecule Discovery Center, University of California, San Francisco, San Francisco, CA, USA.

† These authors contributed equally to this publication.
* To whom correspondence should be addressed.

The ubiquitin system regulates the majority of cellular processes in eukaryotes. Ubiquitin is ligated to substrate proteins as monomers or chains and the topology of ubiquitin modifications regulates substrate interactions with specific proteins. Thus ubiquitination directs a variety of substrate fates including proteasomal degradation[1]. Deubiquitinase enzymes cleave ubiquitin from substrates and are implicated in disease[2]; for example ubiquitin-specific protease-7 (USP7) regulates stability of the p53 tumor suppressor and other proteins critical for tumor cell survival[3]. However, developing selective deubiquitinase inhibitors has been challenging[4] and no co-crystal structures have been solved with small molecule deubiquitinase inhibitors. Here, using nuclear magnetic resonance (NMR)-based screening and structure-based drug design, we describe the development of selective USP7 inhibitors GNE-6640 and GNE-6776. These compounds induce hematopoietic and solid tumor cell death and enhance cytotoxicity with chemotherapeutic agents and targeted compounds, including PIM kinase inhibitors. Structural studies reveal that GNE-6640 and GNE-6776 non-covalently target an interface between the thumb, fingers, and palm regions of the USP7 catalytic domain, 12Å distant from the catalytic cysteine. The compounds attenuate ubiquitin binding and thus inhibit USP7 deubiquitinase activity. Within the USP7 thumb region, GNE-6640 and GNE-6776 interact with acidic residues that mediate H-bond interactions with the ubiquitin Lys-48 side-chain[5], suggesting that USP7 preferentially interacts with and cleaves ubiquitin moieties having free Lys-48 side-chains. We investigated this idea by engineering di-ubiquitin chains containing differential proximal and distal isotopic labels and measuring USP7 binding via NMR, a study that substantiated our hypothesis. This preferential binding significantly protracted the depolymerization kinetics of Lys-48-linked

ubiquitin chains relative to Lys-63-linked chains. In summary, engineering compounds that inhibit USP7 activity by attenuating ubiquitin binding suggests opportunities for developing other deubiquitinase inhibitors and may be a strategy more broadly applicable to inhibiting proteins that require ubiquitin binding for full functional activity.

USP7 is a genetically-validated deubiquitinase for MDM2, a ubiquitin ligase that promotes degradation of the tumor suppressor p53. Depletion of USP7 enhances ubiquitination of USP7 substrates, including MDM2. Subsequent MDM2 proteasomal degradation stabilizes the p53 tumor suppressor, thus promoting cell cycle arrest and apoptosis. USP7 inhibition enhances ubiquitination of additional USP7 substrates that also contribute to cell growth inhibition[3,6,7] (Fig. 1a). To identify USP7 inhibitors, we established a high throughput activity-based screening (HTS) cascade, consisting of a primary deubiquitinase assay followed by a series of counter-screen assays and selection filters (Supplementary Information Fig. 1a, 1b). In parallel we developed an NMR fragment screening cascade to identify fragments that bind the USP7 catalytic domain at discrete sites (Supplementary Information Fig. 1c). A hit-to-lead selection cascade was established to guide optimization of HTS and NMR hits into more potent USP7 inhibitors (Fig. 1b). At the apex of this cascade, deubiquitinase selectivity assays and structural biology directed medicinal chemistry design. Additionally, assays to measure total- and ubiquitinated-MDM2 levels were developed as proximal indicators of USP7 cellular activity, followed by cellular viability assays using inactive compounds from the same chemical series as negative controls (Fig. 1a, 1b).

Progression of HTS and NMR screening hits through the hit-to-lead selection cascade identified a series of fragment compounds as having the most favorable properties for on-target efficacy (Supplementary Information Figs. 2-5). Optimization of the parent compound GNE-2916 yielded GNE-6640 and GNE-6776. These USP7 inhibitors enhanced endogenous MDM2 ubiquitination while achieving selectivity over USP47, the most USP7-homologous deubiquitinase, and USP5, a highly-expressed and active deubiquitinase[8] (Extended Data Fig. 1a). Both compounds accelerated endogenous MDM2 turnover, that was apparent in combination with cycloheximide, a necessary co-treatment since MDM2 is a p53 target gene (Fig. 1a, 1c, Extended Data Fig. 1b). Accordingly, both GNE-6640 and GNE-6776, but not control compounds GNE-2118 and GNE-6641, promoted endogenous MDM2 ubiquitination with Lys-48 (K48)-linked polyubiquitin chains that direct proteins for proteasomal degradation[9] (Fig. 1d, Extended Data Fig. 1c). The active USP7 inhibitors stabilized p53 and upregulated p21 in wild-type cells, but not in p53-null cells, and did not further elevate p53 and p21 expression in USP7-null cells (Fig. 1e, Extended data Fig. 1b). These data were corroborated by a significant dose-dependent decrease in viability of wild-type, but not of USP7-null cells, treated with either inhibitor at concentrations ≤15μM (Extended data Fig. 1d). GNE-6640 also increased p53 and p21 levels and decreased viability to a greater extent in breast adenocarcinoma and osteosarcoma cell lines compared to normal primary breast epithelial cells or osteocytes (Extended Data Fig. 2a, 2b). Thus the collective data indicate that GNE-6640 and GNE-6776 regulate cellular USP7/MDM2/p53 signaling pathways.

A study by Ritorto et al. revealed that previously-reported deubiquitinase inhibitors are neither potent nor sufficiently selective when evaluated at 1-10μM in a MALDI-TOF mass spectrometry-based assay[4]. In this assay, GNE-6640 and GNE-6776 inhibited USP7 cleavage of di-ubiquitin chains and were highly selective against 36 other recombinant

deubiquitinases at 10μM (Fig. 2a, Extended Data Fig. 3a). GNE-6776 remained selective even at 100μM, a >6-fold higher concentration than used in cellular assays, whereas control compound GNE-6641 had negligible activity at either concentration (Fig. 2a, Extended Data Fig. 3b). Since deubiquitinase activities are regulated by numerous cellular mechanisms[10] we developed an assay utilizing activity-based probes to evaluate USP7 inhibitor effects on endogenous deubiquitinases (Heideker et al., manuscript in preparation). Both GNE-6640 and GNE-6776 significantly inhibited USP7 at 15μM (p=0.0503 and 0.0058, respectively; adjusted by the Benjamini-Hochberg False Discovery Rate method) while remaining selective against 44 to 47 other endogenous detected deubiquitinases (Fig. 2b, Extended Data Fig. 3c, 3d). Thus GNE-6640 and GNE-6776 are highly selective USP7 inhibitors against both recombinant and endogenous cellular deubiquitinases. Given these favorable features, we investigated whether the compounds could be utilized in animal studies. Pharmacodynamic and pharmacokinetic studies indicated that GNE-6776 is an orally bioavailable compound that promotes on-target pathway modulation in human hematologic and solid tumor xenografts (Supplementary Information Fig. 6a-e). Although efficacious drug exposure was only transiently achieved, GNE-6776 caused modest, though significant, EOL-1 xenograft growth delay (Supplementary Information Fig. 6f). Thus development of USP7 inhibitors having improved drug-like properties will be required in order to demonstrate the full value of USP7 inhibition *in vivo*.

USP7 and MDM2 regulate proteins other than p53[3,6,11-16] (Fig. 1a). Having confirmed the selectivity and on-target cellular potency of GNE-6640 and GNE-6776, we screened 441 cell lines from 15 tumor indications in three-day viability assays to enable an unbiased analysis of cellular factors that contribute to USP7 inhibitor sensitivity (Extended Data Fig. 4a). GNE-6641 was inactive in all cell lines whereas GNE-6640 decreased viability of 108 cell lines with IC50 ≤10μM (data not shown). Next, we selected a subset of 181 cell lines from six tumor indications that displayed a range of sensitivities for confirmation and more extensive analysis (Extended Data Fig. 4a). GNE-6641 remained inactive in a five-day viability assay, whereas GNE-6640 and GNE-6776 decreased viability of 54 and six cell lines, respectively, with IC50 ≤10μM (Extended Data Fig. 4b). Five-day viability data and exome sequencing analysis were integrated to identify potential features associated with compound sensitivity. Acute myeloid leukemia (AML) cell lines were significantly associated with increased sensitivity to GNE-6640 (Extended Data Fig. 4c, 4e). Loss-of-function mutant TP53 lines were in general less responsive than wild-type TP53 lines, but the trend was not statistically significant (Extended Data Fig. 4d). Indeed, the TP53 R175 hotspot was associated with increased sensitivity while the Y220 hotspot was associated with decreased sensitivity (Extended Data Fig. 4c, 4e). To further investigate the mechanisms by which TP53 status regulates viability in response to USP7 inhibition, we imaged cellular proliferation and caspase activation in TP53 wild-type and TP53-null cell lines treated with USP7 inhibitors. Both compounds decreased proliferation and activated caspases in TP53 wild-type cell lines. These effects were reduced, but not abolished, in the evaluated TP53-null cell lines (Extended Data Fig. 5a-5c). Combining GNE-6640 with doxorubicin or cisplatin, DNA-damaging agents that activate the p53 response[17], enhanced the efficacy of USP7 inhibitors (Extended Data Fig. 6a, 6b). Thus immunoblot analysis and cellular viability studies indicate that the MDM2/p53 pathway is a mediator of GNE-6640 and GNE-6776 cellular activity but is not the only determinant. These findings corroborate genetic studies crossing *TP53*$^{-/-}$ mice with *USP7*$^{-/-}$ mice, that afford only a partial rescue of the *USP7*$^{-/-}$ embryonic lethality[7].

The enhanced efficacy of GNE-6640 in combination with DNA damaging agents suggested a strategy to identify cellular signaling pathways that intersect with USP7 inhibition. We therefore screened the EOL-1 AML cell line using a panel of 589 chemotherapeutic and targeted agents in combination with GNE-6640 or GNE-6776. Top hits for both compounds included inhibitors of PI3K and PIM kinases (Extended Data Fig. 6c, 6d, Supplementary Information Table 1). PIM kinases are reported to share substrates with PI3K pathway kinases including TSC1/2 and the pro-apoptotic Bcl-2 family member Bad (Extended Data Fig. 7d)[18]. Given the role of PIM kinases, particularly PIM2, in hematological malignancies[18], we confirmed combination efficacy with Bliss score analysis using PIM inhibitors GDC-0339[19] or GDC-0570 (Fig. 2c, Extended Data Figures 7a-7c). Both Pim inhibitors share structural similarities and are highly potent against Pim kinases while remaining selective against other kinases (manuscripts in preparation). We also investigated cellular mechanisms of combination efficacy. Whereas GNE-6776 promoted a modest elevation of p21 levels in EOL-1 and MV-4-11 cell lines, PIM2 levels were reduced. The GNE-6776-induced decrease in PIM2 protein level was rescued by inhibition of the ubiquitin-activating enzyme UAE1, that globally blocks ubiquitination of cellular proteins, and the proteasome inhibitor bortezomib, that inhibits degradation of ubiquitinated proteins. PIM2 levels were also lower in USP7 null cells and were rescued by proteasome inhibition, collectively indicating that USP7 stabilizes PIM2 levels via a ubiquitin/proteasome-dependent mechanism (Extended Data Fig. 7e, 7f). Analysis of a treatment time course with GNE-6776, the PIM inhibitor GDC-0570, or the two compounds in combination revealed that, as expected, GDC-0570 reduced phosphorylation of the S6 ribosomal protein and of Bad, and decreased expression of short-lived proteins including p21 and the pro-survival Bcl-2 family member Mcl-1. GNE-6776 decreased PIM2 levels over time and remarkably, also reduced S6 and Bad phosphorylation and total Mcl-1 levels, similar to PIM inhibitor treatment. In combination, GNE-6776 and GDC-0570 treatments enhanced the apoptosis indicators cleaved PARP and cleaved caspase-3, confirming the cell viability combination data (Fig. 2d). We therefore hypothesized that USP7 regulates PIM2 ubiquitination: upon USP7 inhibition, ubiquitin modifications accumulate on PIM2 that promote PIM2 proteasomal degradation. In support of this idea, endogenous USP7 associates with PIM2 in cellular lysates and GNE-6776 treatment enhances PIM2 ubiquitination (Fig. 2e, Extended Data Fig. 7g). Furthermore recombinant USP7, but not a C223S catalytic mutant, deubiquitinated PIM2 *in vitro* (Extended Data Fig. 7h). Our data thus reveal a previously undescribed juncture between USP7 deubiquitinase activity and PIM signaling. This finding expands the repertoire of USP7-regulated oncogenic signaling pathways beyond MDM2/p53 signaling and merits further investigation in additional cell lines and indications.

Given the selectivity and efficacy of GNE-6640 and GNE-6776, it was of interest to understand the precise molecular mechanisms by which these tool compounds inhibit USP7. Enzymatic studies utilizing a ubiquitin-AMC substrate verified that the compounds inhibit intrinsic USP7 catalytic activity (k_{cat}) and notably, attenuate substrate binding (K_m) (Extended Data Fig. 8a). NMR analysis confirmed that USP7 inhibitors weaken USP7 interactions with native ubiquitin (Fig. 3a, Extended Data Fig. 8b, 8c). Analysis of co-crystal structures revealed that both compounds bind into a pocket approximately 12Å from the catalytic triad that is located at the interface of the palm, fingers, and thumb sub-domains (Fig. 3b, Extended Data Fig. 7d, 7e). The phenol-aminopyridine moieties of both compounds make similar interactions with their phenol rings buried in the pocket and hydroxyl groups engaged in H-bond interactions with H403, which is located at the tip of the β8 sheet. The 2-aminopyridine is directed out of

the pocket towards solvent with the 2-amino group engaging D349, which is located on the loop connecting β2 and β3. The two different chemical moieties at the 5-position of the aminopyridine, namely the pyridine-carboxamide (GNE-6776) and indazole (GNE-6440), are largely solvent exposed and lie between the α5 and α6 helices. The NH atoms of the carboxamide of GNE-6776 engages USP7-D305 on the α5 helix and the indazole of GNE-6640 makes extensive van der Waals contacts in the crevice between α5 and α6 (Fig. 3c, Extended Data Fig. 7d). Comparison of the co-crystal structures of GNE-6776 and GNE-6640 with the USP7-ubiquitin complex structure (PDB 1NBF) revealed that the compounds likely sterically inhibit ubiquitin binding and prevent the transition of the USP7 α5 helix to the active conformation[5] (Fig. 3c, Extended Data Fig. 7d). However, given the extensive interactions between USP7 and ubiquitin[5], we investigated the functional significance of residues on the α5 helix of USP7 that interact with K48 in the USP7-ubiqutin structure. The K48A mutant of ubiquitin has reduced affinity for USP7 (Extended Data Fig. 9a). Similarly, recombinant D305A/E308A USP7 catalytic domain failed to cleave tetra-K48- or tetra-K63-linked polyubiquitin conjugated to a TAMRA-labeled peptide[20] and enzymatic analysis revealed this apparent inactivity was due to a significantly increased K_m of 1120 μM (Extended Data Fig. 9b, 9c). Importantly, expression of a D305A/E308A USP7 mutant increased K48-linked polyubiquitination on endogenous MDM2, similarly to expression of catalytically-inactive USP7 (Extended Data Fig. 9d), thus demonstrating the importance of ubiquitin-K48 and USP7-D305/E308 residues for USP7/ubiquitin binding and substrate deubiquitination.

The GNE-6640 and GNE-6776 USP7-interaction modes independently implicated the ubiquitin-K48 side-chain as a key contributor to USP7 binding, suggesting that USP7 preferentially associates with ubiquitin moieties having free K48 side-chains. Mechanistically this could direct USP7 to distal subunits of K48-linked polyubiquitin, resulting in sequential distal-to-proximal chain depolymerization (exo-cleavage). However in non-K48-linked chains, USP7 may bind and cleave proximal to any ubiquitin moieties having accessible K48 side-chains (exo-, endo-, or base-cleavage) (Extended Data Fig. 10a). In order to evaluate this idea, it was necessary to engineer K48- and K63-linked di-ubiquitin with differentially-labeled distal and proximal ubiquitin subunits using spectroscopically distinct amino acids. To select the ubiquitin residues for isotopic labeling, we evaluated the chemical shift perturbations induced by USP7 binding (Extended Data Fig. 10b). The cluster of residues L8, T12, T14, and L15 broaden upon USP7 binding and are well-resolved in the spectra, thus Leu residues were isotopically labeled with ^{15}N in the proximal di-ubiquitin subunit and Thr residues were ^{15}N-labeled in the distal subunit (Fig. 4a, Extended Data Fig. 10b, 10c). NMR analysis of these differentially-labeled di-ubiquitins showed preferential USP7 binding to the distal subunit of K48-linked di-ubiquitin while USP7 binding was similar to both proximal and distal K63-linked di-ubiquitin subunits (Fig. 4a, top panels, Extended Data Fig. 10c). Importantly, both D305A USP7 and D305A/E308A USP7 mutants demonstrated similar binding to both proximal and distal K48-linked di-ubiquitin subunits (Fig. 4a, bottom panels). Based on these data, we evaluated whether differential di-ubiquitin binding impacts the dynamics of ubiquitin chain depolymerization by USP7. A previous study reported that USP7 cleaves K6-, K11-, K33-, K48-, and K63-linked di-ubiquitins with similar k_{cat} and K_m values, indicating that different isopeptide bond conformations do not significantly impact cleavage efficiency[21]. Remarkably, our studies reveal that USP7 sequentially depolymerizes tetra-K48-linked ubiquitin from the distal residue (exo-cleavage) whereas tetra-K63-ubiquitin chains are cleaved via a combination of exo-, endo-, or base-cleavage, as nearly complete depolymerization occurs within four

minutes (Fig. 4b, Extended Data Fig. 10d, 10e). Thus USP7 preferentially binds ubiquitin moieties that have a free K48 side-chain, resulting in protracted and sequential depolymerization of K48-linked polyubiquitin.

Herein we describe the development and characterization of GNE-6640 and GNE-6776, *bona fide* USP7 inhibitors having clear selectivity against other deubiquitinases and a structurally-defined, unique mechanism of action. Establishing a compound selection cascade of biochemical, biophysical, and cellular assays using inactive control compounds for cross-validation was critical for selecting on-target inhibitors and guiding their optimization. Additionally, combination studies revealed a previously undescribed intersection between USP7 deubiquitinase activity and PIM kinases in regulating cell viability. Co-crystal structures of GNE-6640 and GNE-6776 demonstrate that both compounds specifically inhibit USP7 deubiquitinase activity by binding into a unique pocket at the interface of USP7 palm, fingers, and thumb sub-domains and sterically block ubiquitin binding. These inhibitor binding modes also pointed to the potential importance of the complementary charged interactions between USP7-D305/E308 and ubiquitin-K48, which we confirmed via mutational analysis. Notably D305G has been identified as a somatic loss-of-function mutant in acute lymphoblastic leukemia patients[22]. Furthermore, NMR analysis of USP7 binding to native monoubiquitin and differentially-labeled di-ubiquitins revealed that USP7 preferentially interacts with ubiquitin moieties having free K48 side-chains. Thus, although USP7 cleaves most ubiquitin-isopeptide linkages non-selectively[21], USP7 chain depolymerization is markedly protracted for K48-linked polyubiquitin relative to K63-linked polyubiquitin. It has been proposed that the inefficiency of some deubiquitinases to depolymerize longer substrate-conjugated K48-linked chains enables a threshold for proteasome-targeting polyubiquitination[23]; our studies substantiate this idea and provide a biophysical mechanism. Numerous proteins including other deubiquitinases, ubiquitin ligases, DNA repair and endocytosis machinery, and epigenetic regulators are functionally dependent on ubiquitin binding[24]. Developing selective inhibitors that attenuate ubiquitin binding, instead of directly targeting active-site residues, is an effective strategy for USP7 inhibition. Our studies demonstrate the feasibility of this approach, that may have broader applications for inhibiting other classes of ubiquitin-binding proteins.

Acknowledgements

The authors thank Wayne Fairbrother, Lauren Frick, Sharon Fong, Tom Hunsaker, Cynthia Lam, Bianca Liederer, Mark Merchant, Jim Nonomiya, Jing Peng, Thin Pham, Linda Rangell, Richard A. Rodriguez, Udi Segal, Raymond Tong, Leslie Wang, and the Genentech Protein Expression group, Cell Central group, gCSI group, and the Boston Biochem, Inc. team for reagents, helpful discussions, and collaborations. DRA and MT are supported by the UK Medical Reseach Council (grant number MC_UU_12016/2).

References

1 Komander, D. & Rape, M. The ubiquitin code. *Annual review of biochemistry* **81**, 203-229, doi:10.1146/annurev-biochem-060310-170328 (2012).

2 Clague, M. J., Heride, C. & Urbe, S. The demographics of the ubiquitin system. *Trends Cell Biol* **25**, 417-426, doi:10.1016/j.tcb.2015.03.002 (2015).

3 Nicholson, B. & Suresh Kumar, K. G. The multifaceted roles of USP7: new therapeutic opportunities. *Cell Biochem Biophys* **60**, 61-68, doi:10.1007/s12013-011-9185-5 (2011).

4 Ritorto, M. S. *et al.* Screening of DUB activity and specificity by MALDI-TOF mass spectrometry. *Nature communications* **5**, 4763, doi:10.1038/ncomms5763 (2014).

5 Hu, M. *et al.* Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde. *Cell* **111**, 1041-1054 (2002).

6 Cummins, J. M. *et al.* Tumour suppression: disruption of HAUSP gene stabilizes p53. *Nature* **428**, 1 p following 486, doi:10.1038/nature02501 (2004).

7 Kon, N. *et al.* Inactivation of HAUSP in vivo modulates p53 function. *Oncogene* **29**, 1270-1279.

8 Clague, M. J. *et al.* Deubiquitylases from genes to organism. *Physiol Rev* **93**, 1289-1315, doi:10.1152/physrev.00002.2013 (2013).

9 Deshaies, R. J. Proteotoxic crisis, the ubiquitin-proteasome system, and cancer therapy. *BMC biology* **12**, 94, doi:10.1186/s12915-014-0094-0 (2014).

10 Heideker, J. & Wertz, I. E. DUBs, the regulation of cell identity and disease. *The Biochemical journal* **467**, 191 (2015).

11 Fahraeus, R. & Olivares-Illana, V. MDM2's social network. *Oncogene* **33**, 4365-4376, doi:10.1038/onc.2013.410 (2014).

12 Song, M. S. *et al.* The deubiquitinylation and localization of PTEN are regulated by a HAUSP-PML network. *Nature* **455**, 813-817, doi:10.1038/nature07290 (2008).

13 Pozhidaeva, A. K. *et al.* Structural Characterization of Interaction between Human Ubiquitin-specific Protease 7 and Immediate-Early Protein ICP0 of Herpes Simplex Virus-1. *The Journal of biological chemistry* **290**, 22907-22918, doi:10.1074/jbc.M115.664805 (2015).

14 Cheng, J. *et al.* Molecular mechanism for USP7-mediated DNMT1 stabilization by acetylation. *Nature communications* **6**, 7023, doi:10.1038/ncomms8023 (2015).

15 Tavana, O. *et al.* HAUSP deubiquitinates and stabilizes N-Myc in neuroblastoma. *Nat Med* **22**, 1180-1186, doi:10.1038/nm.4180 (2016).

16 Wang, L. *et al.* Ubiquitin-specific Protease-7 Inhibition Impairs Tip60-dependent Foxp3+ T-regulatory Cell Function and Promotes Antitumor Immunity. *EBioMedicine* **13**, 99-112, doi:10.1016/j.ebiom.2016.10.018 (2016).

17 Williams, A. B. & Schumacher, B. p53 in the DNA-Damage-Repair Process. *Cold Spring Harb Perspect Med* **6**, doi:10.1101/cshperspect.a026070 (2016).

18 Nawijn, M. C., Alendar, A. & Berns, A. For better or for worse: the role of Pim oncogenes in tumorigenesis. *Nat Rev Cancer* **11**, 23-34, doi:10.1038/nrc2986 (2011).
19 Hodges, A. J. M., M.; Sharpe, A.; Sun, M.; Wang, X.; Tsui, V. H. Pyrazol-4-yl-heterocyclyl-carboxamide compounds as Pim kinase inhibitors and their preparation. PCT Int. Appl. WO2013045461 (2013).
20 Geurink, P. P., El Oualid, F., Jonker, A., Hameed, D. S. & Ovaa, H. A general chemical ligation approach towards isopeptide-linked ubiquitin and ubiquitin-like assay reagents. *Chembiochem : a European journal of chemical biology* **13**, 293-297, doi:10.1002/cbic.201100706 (2012).
21 Faesen, A. C. *et al.* The differential modulation of USP activity by internal regulatory domains, interactors and eight ubiquitin chain types. *Chem Biol* **18**, 1550-1561, doi:10.1016/j.chembiol.2011.10.017 (2011).
22 Huether, R. *et al.* The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. *Nature communications* **5**, 3630, doi:10.1038/ncomms4630 (2014).
23 Schaefer, J. B. & Morgan, D. O. Protein-linked ubiquitin chain structure restricts activity of deubiquitinating enzymes. *The Journal of biological chemistry* **286**, 45186-45196, doi:10.1074/jbc.M111.310094 (2011).
24 Husnjak, K. & Dikic, I. Ubiquitin-binding proteins: decoders of ubiquitin-mediated cellular functions. *Annual review of biochemistry* **81**, 291-322, doi:10.1146/annurev-biochem-051810-094654 (2012).
25 Reverdy, C. *et al.* Discovery of specific inhibitors of human USP7/HAUSP deubiquitinating enzyme. *Chem Biol* **19**, 467-477, doi:10.1016/j.chembiol.2012.02.007 (2012).
26 Yu, M. *et al.* A resource for cell line authentication, annotation and quality control. *Nature* **520**, 307-311, doi:10.1038/nature14397 (2015).
27 Wertz, I. E. *et al.* Phosphorylation and linear ubiquitin direct A20 inhibition of inflammation. *Nature* **528**, 370-375, doi:10.1038/nature16165 (2015).
28 Haverty, P. M. *et al.* Reproducible pharmacogenomic profiling of cancer cell line panels. *Nature* **533**, 333-337, doi:10.1038/nature17987 (2016).
29 Bett, J. S. *et al.* Ubiquitin C-terminal hydrolases cleave isopeptide- and peptide-linked ubiquitin from structured proteins but do not edit ubiquitin homopolymers. *The Biochemical journal* **466**, 489-498, doi:10.1042/BJ20141349 (2015).
30 Iwasaki, T. *et al.* Dissection of hydrogen bond interaction network around an iron-sulfur cluster by site-specific isotope labeling of hyperthermophilic archaeal Rieske-type ferredoxin. *J Am Chem Soc* **134**, 19731-19738, doi:10.1021/ja308049u (2012).

Figure Legends

Figure 1: Identification and characterization of USP7 inhibitors. a. Schematic of USP7-regulated signaling. USP7 inhibition increases substrate ubiquitination, leading to proteasome-mediated substrate degradation. Increased ubiquitination and subsequent degradation of the MDM2 ubiquitin ligase stabilizes the p53 tumor suppressor, resulting in cell cycle arrest or cell death. MDM2 is also a p53-target gene. **b.** A schematic of the USP7 inhibitor hit-to-lead selection cascade. See text for more details. **c.** Left panels: western blot analysis of MDM2 turnover upon USP7 inhibition. MCF7 cells were treated with cycloheximide (CHX) for the indicated times with DMSO vehicle or the indicated compounds. See Extended Data Fig. 1b for corresponding USP7, p53, and p21 western blots. Right panel: graph showing actin-normalized MDM2 levels from western blots of three experiments, error bars indicate SD. **d.** Analysis of endogenous MDM2 polyubiquitinated with K48-linked chains. Top panel: denatured lysates from HCT-116 cells treated for 8 hours with the indicated compounds were immunoprecipitated with a K48 polyubiquitin linkage-specific antibody and immunocomplexes were western blotted with an anti-MDM2 antibody. Western blot analysis of whole cell lysates for the indicated proteins are shown below. **e.** Western blot analysis of p53-pathway proteins in wild-type, USP7 null, or p53-null HCT116 cells after 24-hour treatment with 15μM of the indicated USP7 inhibitors or inactive controls. At least three experimental replicates were performed for panels **c**. - **e**.

Figure 2: Selectivity of USP7 inhibitors and mechanism of synergy with PIM kinase inhibition. a. An activity heatmap of a panel of deubiquitinases incubated with USP7 inhibitors. Deubiquitinases at the indicated concentrations were treated with 10μM of the indicated USP7 inhibitor compounds, incubated with di-ubiquitin substrates, and analyzed by MALDI-TOF mass spectrometry. Detection of monoubiquitin was used to quantitate deubiquitinase activity, shown in a gradient of white (0% inhibition) to dark red (100% inhibition). **b.** Volcano plots summarizing activity-based profiling data of endogenous deubiquitinases treated with USP7 inhibitors. Benjamini-Hochberg False Discovery Rate-adjusted *p*-values are plotted vs. normalized $\log_2$ fold change in deubiquitinase activity in 293T cells treated in experimental triplicates (left graph) or quadruplicates (right graph) with the indicated USP7 inhibitors compared to the indicated controls. **c.** Bliss analysis of 9x9 dose response matrix with PIM inhibitor GDC-0570 and GNE-6776 in EOL-1 cells. Left panel: curve fitted viability values at each dose across the matrix. Zero represents no effect whereas 100 indicates complete loss of viability. Right panel: the difference in observed versus predicted values using the Bliss independence model. Positive values indicate a greater than predicted decrease in viability. **d.** Time course study of EOL-1 cells treated with 2μM GNE-6776, 0.02μM GDC-0570, or a combination of the two compounds. Cells were collected at the indicated time points and the indicated proteins were examined by immunoblotting cell lysates. ** indicates uncleaved PARP; * indicates cleaved PARP. **e.** Analysis of endogenous PIM2 polyubiquitinated with K48-linked chains. Top panel: denatured lysates from EOL-1 cells treated for the indicated times with GNE-6776 were immunoprecipitated with a K48 polyubiquitin linkage-specific antibody and immunocomplexes were western blotted with an anti-PIM2 antibody. Western blot analysis of whole cell lysates for the indicated proteins are shown below.

Figure 3: USP7 inhibitors compete with ubiquitin binding to USP7. a. Overlay of a region of the 2D $^{1}H/^{15}N$ transverse relaxation optimized spectroscopy (TROSY) spectrum of the USP7 catalytic domain (orange) highlighting changes induced by binding of

ubiquitin in the absence (left, blue) and presence of GNE-6776 (middle, dark grey). The right panel shows the effect of GNE-6776 on labeled USP7 in the absence of ubiquitin in light gray. Individual peaks stemming from residues E371 and Q287 are highlighted. Three experimental replicates were performed. **b.** Left panel: crystal structure of USP7 catalytic domain in complex with GNE-6776. The catalytic domain is shown as an orange cartoon, the ligand as yellow sticks, and the catalytic triad residues are shown as green sticks with individual atoms colored following CPK color convention. Right panel: GNE-6776 compound structure. **c.** Left panel: depiction of the USP7 catalytic domain/ubiquitin complex PDB 1NBF (green and cyan cartoons, respectively) with the K48 residue side chain in ubiquitin and the acidic patch D305 and E308 in USP7 represented as sticks. Binding interactions are shown as gray dashed lines. Right panel: the structure of USP7/GNE-6776 (orange cartoon and residue side chains and yellow sticks, respectively) and the structure of ubiquitin (cyan) modeled in from PDB 1NBF. GNE-6776 sterically blocks the binding of ubiquitin and prevents the hydrogen bond interaction between ubiquitin-K48 and USP7-D305. The E308 side chain is also oriented away from ubiquitin-K48.

Figure 4: USP7 preferentially binds and cleaves ubiquitin moieties with free K48 side-chains. a. Schematic diagrams of differentially labeled di-ubiquitins (also see Extended Data Fig. 10b, 10c) and representative 2D $^{1}H/^{15}N$ SOFAST spectra overlays, and 1D traces extracted from the ^{1}H dimension at the peak maximum. In both K48- and K63-linked di-ubiquitin schematics, the distal (dist.) ubiquitin has isotopically-labeled ^{15}N-Thr residues and the proximal (prox.) ubiquitin has ^{15}N-Leu residues, marked with asterisks. Top left panel: K48-linked di-ubiquitin spectra and 1D traces. Di-ubiquitin ^{15}N-Thr or ^{15}N-Leu peaks are shown in green; the same peaks after addition of unlabeled USP7-C223A catalytic domain are overlaid in orange. Top right panel: K63-linked di-ubiquitin spectra and 1D traces. Di-ubiquitin ^{15}N-Thr or ^{15}N-Leu peaks are shown in purple; the same peaks after addition of unlabeled USP7-C223A catalytic domain are overlaid in orange. Bottom panels: the SOFAST experiments with K48-linked di-ubiquitin ^{15}N-Thr or ^{15}N-Leu peaks shown in green and the effect of adding the USP7 catalytic domain double mutant C223A/D305A (left) or triple mutant C223A/D305A/E308A (right) overlaid in orange. The concentration ratios for the top two experiments were 1:1 at 70μM ubiquitin, the bottom left was 1:1.5 at 65μM ubiquitin and the bottom right was 1:2 at 60μM ubiquitin. **b.** Time course analysis of peptide-conjugated tetra-ubiquitin chains cleaved by full-length USP7 (above) and corresponding densitometry plots (below). Left panels: time course of USP7-mediated depolymerization of K48-linked tetra-ubiquitin conjugated to a TAMRA-labeled peptide. Right panels: time course of USP7-mediated depolymerization of K63-linked tetra-ubiquitin conjugated to a TAMRA-labeled peptide. At least three experimental replicates were performed.

Extended Data Figure Legends

Extended Data Figure 1: Deubiquitinase inhibition and cellular activity of optimized fragment compounds.
a. A table summarizing deubiquitinase biochemical assay data and ubiquitin-MDM2 assay data from optimized fragment compounds and inactive controls. Fragment compound structures are shown below. **b.** Western blot analysis of USP7, p53, p21 levels from the cycloheximide (CHX)-chase study of MDM2 turnover shown in Fig. 1c. **c.** Analysis of endogenous MDM2 polyubiquitinated with K48-linked chains. Top panel: denatured lysates from MCF7 cells treated for 8 hours with the indicated compounds were immunoprecipitated with a K48 polyubiquitin linkage-specific antibody and immunocomplexes were western blotted with an anti-MDM2 antibody. Western blot analysis of whole cell lysates for the indicated proteins are shown below. **d.** Cell viability of wild-type and USP7-null HCT116 colon adenocarcinoma cells, treated as indicated, and analyzed using the CellTiter-Glo assay. Data normalized to vehicle control are plotted as a function of compound concentration. Two-sided t-tests were used to calculate *p*-values between wild-type HCT-116 and USP7-null cells treated with GNE-6640. 7.5μM $p = 0.01$, 10μM $p = 0.041$, 12.5μM $p = 0.009$, 15μM $p = 0.011$, 20μM $p = 0.017$. **e.** Cell viability of wild-type and USP7-null HCT116 colon adenocarcinoma cells, treated and analyzed as in **d.** with the indicated doses of GNE-6776. 1μM $p = 0.023$, 2.5μM $p = 0.003$, 5μM $p = 0.001$, 7.5μM $p = 0.003$, 10μM $p = 0.007$, 12.5μM $p = 0.001$, 15μM $p = 0.007$, 17.5μM $p = 0.001$, 20μM $p = 0.008$. At least two experimental replicates were performed.

Extended Data Figure 2: Effects of USP7 inhibitors in human primary cells and tumor cell lines.
a. Top panels: representative western blots of p53-pathway proteins and tubulin loading controls in lysates of tumor cell lines and tissue-matched primary cells after 24 hour treatment with USP7 inhibitors. Lower graphs: bands from p53, p21, and tubulin immunoblots were quantified and p53 and p21 expression was normalized to tubulin. The tubulin-normalized p53 or p21 ratio of the DMSO-treated sample in the relevant primary cell line was arbitrarily set to a value of one. The average relative ratios of tubulin-normalized p53 or p21 expression levels are plotted for four (breast cell lines) or three (osteo-cell lines) biological replicate experiments. Asterisks indicate $p<0.05$, two-sided t-test. **b.** Cell viability of tumor cell lines and tissue-matched primary cells, treated as indicated, and analyzed using the CellTiter-Glo assay. Data normalized to vehicle control are plotted as a function of compound concentration. At least three experimental replicates were performed for panels a and b.

Extended Data Figure 3: Selectivity and cellular efficacy of USP7 inhibitors.
a. GNE-6776 dose-dependent inhibition of di-ubiquitin cleavage by USP7 measured by MALDI-TOF and plotted using SigmaPlot v.12.5. **b.** Percent inhibition of the indicated deubiquitinases for cleaving di-ubiquitins after treatment with 100μM of the indicated USP7 inhibitor compounds. Deubiquitinase concentrations and di-ubiquitin substrates are as in Fig. 2a. **c.** Supporting western blots for Fig. 2b, left panel. HEK-293T cell lysates were treated with the indicated USP7 inhibitors (0.1% DMSO control = 0μM compound) and endogenous deubiquitinases were reacted with the HA-ubiquitin-vinylsulfone activity-based probe (HA-Ub-VS). Reacted cell lysates were immunoblotted with the indicated antibodies. * indicates unreacted deubiquitinases, ** indicates probe-reacted deubiquitinases, the arrowhead points to a band identified by anti-HA immunoblotting that runs at the expected molecular weight of USP7 and is diminished in

lysates treated with GNE-6640. **d.** Supporting western blots for Fig. 2b, right panel. HEK-293T cell lysates were treated with the indicated USP7 inhibitors, endogenous deubiquitinases were reacted with the HA-Ub-VS activity-based probe, and reacted cell lysates were immunoblotted with the indicated antibodies. * indicates unreacted deubiquitinases, ** indicates probe-reacted deubiquitinases.

Extended Data Figure 4: Bioinformatics analysis of USP7 inhibitor cell viability screens.

a. Schematic of the cellular viability assay workflow and bioinformatics analysis. The six tumor cell line indications included leukemias, lymphomas, lung carcinomas, and breast, colon, and prostate adenocarcinomas. **b.** Histogram of IC_{50} values of GNE-6640, GNE-6446, and GNE-6641 in 181 cell lines. Mean viability is calculated as the arithmetic average of the fitted viabilities at each tested dose of GNE-6640 or GNE-6446 normalized to the mean viability of GNE-6641. **c.** Univariate analysis of features associated with viability differences. The x-axis represents the fold change (log_2) in normalized mean viability between cell lines present or absent for a given feature. The y-axis represents the nominal *p*-value ($-log_{10}$ scale). Features with q-values less than 0.05 and absolute log_2 fold change greater than 0.1 are colored in red. Features with only absolute log_2 fold change greater than 0.1 are colored in gold. *P*-values were determined using the two-sided Student's t-Test and q-values were determined by correcting resulting *p*-values for multiple hypothesis testing using the Benjamini and Hochberg approach. The size of each point corresponds to the number of cell lines present with the feature. **d, e.** Boxplots of selected features and their respective associations with normalized mean viability. The respective *p*- and q-values are indicated below.

Extended Data Figure 5: Live cell imaging of USP7 inhibitor-treated cells.

Graphs showing cell confluence as a function of time (top rows) and normalized caspase activity (bottom rows) in cells treated with the indicated USP7 inhibitors. **a.** TP53 wild-type or TP53 null HCT116 colon adenocarcinoma cells. **b.** TP53 wild-type MCF7 or TP53 null MDA-MB157 breast adenocarcinoma cells. **c.** TP53 wild-type U2OS cells or TP53 null SaOS osteosarcoma cells. At least three experimental replicates were performed for all experiments.

Extended Data Figure 6: Viability of USP7 inhibitor-treated cells in combination with chemotherapeutic and targeted agents.

Graphs showing cell confluence as a function of time (top rows) and normalized caspase activity (bottom rows) in cells treated with the indicated USP7 inhibitors and/or chemotherapeutics. **a.** MCF7 breast adenocarcinoma cells treated with GNE-6640 or doxorubicin alone or in combination. **b.** U2OS osteosarcoma cells treated with GNE-6640 or cisplatin alone or in combination. At least three experimental replicates were performed for all experiments. **c.** A pie chart illustrating the distribution of compound classes in the Genentech Chemical Genomics Compound library, comprising 589 compounds. NHR = nuclear hormone receptor, GEF = guanine nucleotide exchange factor, DNA = DNA damaging agent. **d.** Bar plot visualizing the -log10 transformed *p*-value from the Wilcoxon rank sum test evaluating the enrichment of a given compound target over all concentrations of USP7 inhibitors vs. DMSO experiments in EOL-1 cells. Only compound targets with 3 or more compounds in the screen were visualized. Higher values indicate synergy with USP7 inhibitors and were followed up by Bliss analysis. Compounds common to certain signaling pathways including PI3K/PIM, RTK/MAPK, epigenetic regulation, and DNA damage are color-coded as indicated.

Extended Data Figure 7: Mechanism of action studies with USP7 inhibitor and PIM inhibitor combinations. **a.** Compound structure of GDC-0339. **b.** PIM inhibitor viability curves at different fixed doses of GNE-6676 in EOL-1 cells. **c.** Bliss analysis of 9x9 dose response matrix with PIM inhibitor GDC-0339 and GNE-6776 in EOL-1 cells. Left panel: curve-fitted viability values at each dose across the matrix. Zero represents no effect whereas 100 indicates complete loss of viability. Right panel: the difference in observed versus predicted values using the Bliss independence model. Positive values indicate a greater than predicted decrease in viability. **d.** A schematic of the PI3K signaling pathway and regulation by PIM kinases. PIM and AKT kinases regulate Bad and TSC1/2 phosphorylation status. Phospho-proteins highlighted in yellow were profiled in cellular studies shown in Figure 2d. **e.** Immunoblot analysis of cell lysates from the indicated cell lines treated with GNE-6776 (2μM for 18 hours), a UAE1 inhibitor MLN-7243 (5μM for 45 minutes) or the proteasome inhibitor bortezomib (5μM for 45 minutes). **f.** Immunoblot analysis of cell lysates from the indicated cell lines, either untreated or treated with the proteasome inhibitor bortezomib (5μM for 45 minutes). **g.** Immunoprecipitation of cellular lysates using an anti-USP7 antibody or an isotype-matched control antibody indicates a specific interaction between endogenous USP7 and PIM2. **h.** Wild-type recombinant USP7, but not a catalytically inactive USP7 mutant, deubiquitinates endogenous polyubiquitinated PIM2 that was immunoprecipitated from proteasome inhibitor-treated MV-4-11 cells.

Extended Data Figure 8: Enzymatic analysis and supporting structural biology data for USP7 inhibitors and USP7.
a. Michaelis-Menten kinetic analysis of USP7 and a series of ubiquitin-AMC substrate titrations with the indicated USP7 inhibitors. Initial rate of substrate hydrolysis was determined using the Magellan software on a Tecan Safire2 plate reader and kinetic parameters were modeled using nonlinear regression analysis with GraphPad Prism software. Standard error was calculated from at least 3 experimental replicates. **b.** Affinity values of ubiquitin binding to USP7 catalytic domain in the absence and presence of USP7 inhibitors. The values were determined by titration of unlabeled ubiquitin to labeled USP7 catalytic domain and the NMR chemical shift changes were fitted as described in the methods. Standard error was calculated from at least 3 experimental replicates. **c.** Overlay of a region of the 2D $^{1}H/^{15}N$ transverse relaxation optimized spectroscopy (TROSY) spectrum of the USP7 catalytic domain (orange) highlighting changes induced by binding of ubiquitin in the absence (left, blue) and presence of GNE-6640 (right, black). Individual peaks stemming from residues E371 and Q387 are highlighted. Three experimental replicates were performed. **d.** Comparison of the crystal structure of USP7 catalytic domain in complex with GNE-6640 (cyan) and GNE-6776 (yellow). The catalytic domain is shown as an orange cartoon and the side chains of the residues in proximity to the inhibitor binding sites are shown as orange sticks. GNE-6640 and GNE-6776 compound structures are indicated above. **e.** Data collection and refinement statistics for GNE-6440 and GNE-6776 crystal structures with the USP7 catalytic domain.

Extended Data Figure 9: Analysis of the functional significance of the interactions between ubiquitin-K48 and USP7-D305, -E308 residues.
a. Titration curves showing the effect of unlabeled wild-type ubiquitin (top) and ubiquitin K48A (bottom) addition to [$^{2}H/^{15}N$] labeled USP7 catalytic domain. The weighted combined $^{1}H/^{15}N$ chemical shift change is plotted against the ubiquitin concentration for 5 well-resolved peaks stemming from E371, Q387, A381, D342 and Y339 residues in the ^{15}N TROSY spectrum. Standard error was calculated from at least three experimental

replicates. **b.** Time course analysis of peptide-conjugated tetra-ubiquitin chains reacted with the USP7 catalytic domain D305A/E308A mutant. **c.** Michaelis-Menten analysis of USP7 catalytic domain D305A/E308A mutant showing the results of three independent experiments. **d.** Evaluation of endogenous MDM2 ubiquitination status upon expression of wild-type, C223A, or D305A/E308A full-length USP7. Top panel: denatured lysates from MCF7 cells transfected with the indicated USP7 expression constructs were immunoprecipitated with a K48 linkage-specific antibody and immunocomplexes were blotted with an MDM2 antibody. Western blot analysis of whole cell lysates for the indicated proteins are shown below. Three experimental replicates were performed.

Extended Data Figure 10: Analysis of differential USP7 binding to K48- and K63-linked poly-ubiquitin and depolymerization kinetics.

a. Schematic diagrams of substrate-bound K48-linked polyubiquitin chains (left) and K63-linked chains (right), and proposed USP7 interactions. Dashed lines indicate residue side-chains and the • symbol indicates an isopeptide bond between the ubiquitin C-terminus and a Lys residue side chain. The proximal ubiquitin subunits (ligated to the substrate) and their K48 or K63 side-chains are indicated with "a" subscripts, the next most distal ubiquitin subunits and side-chains are indicated with "b" subscripts, etc. Preferential USP7 binding to free K48 side-chains would direct USP7 to the distal ubiquitin subunit of K48 polyubiquitin and promote sequential exo-cleavage, whereas USP7 would bind all subunits of K63-polyubiquitin and promote exo-, endo-, and base-cleavage. **b.** The $^{1}H/^{15}N$ SOFAST spectrum of labeled ubiquitin (cyan), superimposed with the spectrum of a 1:1 molar ration of labeled ubiquitin in the presence of unlabeled USP7-catalytic domain (orange). Ubiquitin residues affected by the USP7 interaction, that results in exchange broadening of the residue cross peaks, are labeled with grey or blue text and correspond to the residues depicted in Extended Data Fig. 10c. The K48 and K63 residues are labeled with red font; the L43 and L50 residues labeled with green font do not broaden upon USP7 binding and serve as internal controls. The box highlights the region depicted in Fig. 4a. **c.** Structure depictions by ribbon diagrams of the covalent complex between USP7 catalytic domain (orange) with ubiquitin (cyan); (PDB code 1NBF) [top panel], K48 linked di-ubiquitin (green; PDB code 2KDE) [lower left panel], and K63 linked di-ubiquitin (purple; PDB code 2RR9) [lower right panel]. In all diagrams, highlighted spheres are the residues in ubiquitin that are broadened in the $^{1}H/^{15}N$ SOFAST spectrum of $^{1}H/^{15}N$ labeled protein by addition of unlabeled USP7 catalytic domain (see Extended Data Fig. 10b for more details). Leu or Thr residues colored in blue show well-resolved peaks and were amenable to selective ^{15}N labeling of di-ubiquitin, highlighted with asterisks in the corresponding schematic diagrams. Lysine side chains of K48 and K63 are indicated as sticks in red. **d.** Time course analysis of peptide-conjugated tetra-ubiquitin chains cleaved by the USP7 catalytic domain (above) and corresponding densitometry plots (below). Left panels: time course of USP7 catalytic domain-mediated depolymerization of K48-linked tetra-ubiquitin conjugated to a TAMRA-labeled peptide. Right panels: time course of USP7 catalytic domain-mediated depolymerization of K63-linked tetra-ubiquitin conjugated to a TAMRA-labeled peptide. **e.** A shorter 0 – 7 minute time-course analysis of TAMRA peptide-K63 tetra-ubiquitin conjugate depolymerization by full-length USP7 (top gel) and the corresponding densitometry plot (below). At least three experimental replicates were performed.

Supplementary Information figure legends and summary

Supplementary Information Figure 1: Screening cascades for USP7 inhibitors. **a.** High-throughput activity-based screening cascade to identify USP7 inhibitors. Screening stages are identified in bold print. Numbers of compounds at each stage are provided to the right of each box. Criteria for progression to the next stage are highlighted in italics to the left of each arrow. **b.** Confirmed hits from USP7 screen. Structures and assay results are provided for 5 of the 101 confirmed actives. The compounds were clustered by structural similarity. IC_{50} values and Hill slopes were determined from 10-point dose titrations with n=2. **c.** Fragment NMR screen diagram. Screening stages are identified in bold print. Numbers of compounds at each stage are provided to the right of each box. Criteria for progression to the next stage are highlighted in italics to the left of each arrow. Protein Saturation Transfer Difference (STD) experiments were performed at 283K. Primary USP7 catalytic domain binders were selected based on the signal to noise (S/N) of the respective compound with a cut-off of greater than 5. All primary binders were re-measured as single compounds under otherwise identical conditions and confirmed binders selected having a S/N of greater than 10. Hits were further tested for specific binding to USP7 catalytic domain by measurement of $^{1}H^{15}N$ TROSY protein spectra. Positive hits were defined as compounds that induced chemical shift perturbations. Perturbations were classified by the chemical shift patterns and selected compounds passed onto X-ray soaking experiments.

Supplementary Information Figure 2: Hit-to-lead selection assay data summary for lead USP7 inhibitors.
A table summarizing the hit-to-lead assay results of the lead compounds identified by the high throughput screening and NMR fragment screening campaigns. Compound series are grouped in columns and hit-to-lead selection assay data are listed in the indicated rows. See Fig. 1b and text for more details.

Indole tricyclic compounds including GNE-8735 increased total MDM2 levels and inhibited cathepsin-B, indicating poor selectivity and induction of general cell stress by this chemical series (see also Supplementary Information Fig. 3a and 3b). Indole tricyclics also precipitated caspase-3, although they passed dynamic light scattering (DLS) analysis (see also Supplementary Information Fig. 4a). The peptidomimetic compounds had weaker biochemical potency, poor selectivity, and covalently modified USP7 cysteine (Cys) residues other than the catalytic Cys (data not shown). Given these data, and because optimization of indole tricyclic and peptidomimetic compounds proved challenging, these series were discontinued. The tetrahydroacridine and fragment compounds were relatively potent, selective, and enhanced cellular MDM2 ubiquitination without significantly increasing total MDM2 (see also Supplementary Fig. 3a and 3b). Tetrahydroacridine compounds passed cathepsin-B inhibition assays, demonstrating additional protease selectivity. Neither tetrahydroacridine nor fragment compounds showed evidence of USP7 aggregation in DLS or in NMR studies (see Supplementary Fig. 4a and data not shown). Tetrahydroacridine compounds, including GNE-6831, covalently modified USP7, consistent with a previous report describing a similar series[25] (see also Supplementary Information Fig. 4b).

Supplementary Information Figure 3: Analysis of total- and ubiquitin-MDM2 levels in cells.
a. Cellular MDM2 immunofluorescence studies. HCT-116 human colorectal carcinoma cells were treated with a range of concentrations of the indicated USP7 inhibitors or DMSO vehicle for 24 hours and endogenous MDM2 protein levels were detected by immunofluorescence imaging. Depicted images show cells treated with 10 μM of the indicated compounds or DMSO vehicle control. The graph shows the quantified mean nuclear MDM2 protein levels per cell over a range of concentrations of GNE-8735 and GNE-2916 (error bars represent standard deviation). The half maximal effective concentration (EC_{50}) for the elevation in MDM2 caused by GNE-8735 was 2.9 μM. **b.** Quantitation of total- and ubiquitinated-MDM2 (Ub-MDM2) in USP7 inhibitor-treated cells. SJSA-1 human osteosarcoma cells were treated with a range of concentrations of the indicated USP7 inhibitors or DMSO vehicle control for 24 hours and the level of ubiquitinated-MDM2 and total MDM2 were measured using a multiplexed Mesoscale immunoassay. Representative graphs show the quantified level of either total MDM2 (right column, red), ubiquitinated MDM2 (central column, blue) or the ratio of the ubiquitinated-MDM2 signal and the total MDM2 signal (left column, orange). All data are shown as percentage change in each value, relative to DMSO vehicle-treated samples. The maximal extent of the increase in the ubiquitinated-MDM2/total MDM2 ratio varied between compounds and in most cases did not reach a plateau. In order to compare the potency of the on the increase in the ratio of ubiquitinated-MDM2/total MDM2 between compounds, the top level was set to 100% and was universally applied to calculate a value for the half maximal effective concentration (EC_{50}) of the percent change in this ratio relative to DMSO (shown in the left column). At least two experimental replicates were performed.

Supplementary Information Figure 4: Biophysical analysis of selected USP7 inhibitors.
a. The DLS autocorrelation functions are shown for 100μM Rottlerin (red), full-length USP7 with 100μM GNE-8735 (blue), full-length USP7 with 100μM GNE-2090 (brown) and full-length USP7 with 0.1% DMSO vehicle control (black). The percent aggregate by mass is shown in the table. **b.** USP7 full-length protein was incubated overnight with excess of GNE-6831 and analyzed by LC-MS. Unmodified and covalently modified USP7 are represented in the top and bottom panels, respectively. Three experimental replicates were performed.

Supplementary Information Figure 5: Deubiquitinase inhibition, ubiquitin-MDM2, and cellular viability data for tetrahydroacridine and fragment compounds.
a. A table of active and inactive tetrahydroacridine compounds with structures. Hit-to-lead selection assay data are listed in the indicated rows. See Fig. 1b and text for more details. **b.** A table of active and inactive fragment compounds with structures. Hit-to-lead selection assay data are listed in the indicated rows. See Fig. 1b and text for more details. **c.** Cell viability assays in AMO-1 cells treated as indicated with the tetrahydroacridine compounds (top graph; purple lines are inactive controls and the green line is the active compound) and fragment compounds (bottom graph; red lines are inactive controls and the blue line is the active compound). Data normalized to vehicle control are plotted as a function of compound concentration. **d.** Cell viability assays in KMS12-PE cells treated as indicated with the tetrahydroacridine compounds (top graph; purple lines are inactive controls and the green line is the active compound) and fragment compounds (bottom graph; red lines are inactive controls and the blue line

is the active compound). Data normalized to vehicle control are plotted as a function of compound concentration. At least two experimental replicates were performed.

Tetrahydroacridine compounds GNE-6831 and GNE-2090 decreased viability of KMS12-PE and AMO1 multiple myeloma cell lines but this activity was not differentiated from control compounds GNE-0956, -2143 and -2148. In contrast, the fragment compound GNE-2916 decreased multiple myeloma cell viability significantly more than control compounds GNE-2917, -2931, and -9603. Thus work on tetrahydroacridine series was discontinued and we focused on optimizing the fragment series.

Supplementary Information Figure 6: *in vitro* and *in vivo* drug metabolism and pharmacokinetic (DMPK) profiling, pharmacodynamic effects, and xenograft growth inhibition studies with GNE-6776.

a. *In vitro* pharmacokinetic assessment of USP7 inhibitors. Calculated drug properties are indicated: molecular weight (MW), lipophilicity at pH 7.4 ($LogD_{7.4}$), total polar surface area (tPSA), stability in hepatic microsomes (LM CL_{hep}) or hepatocytes (Hep CL_{hep}) from human/rat/mouse/dog/cyno (h/r/m/d/c) species, percent plasma protein binding (PPB %) and permeability across an MDCK cell monolayer from basolateral to apical (B to A) or apical to basolateral (A to B) directions. **b.** EOL-1 cell line viability in response to GNE-6776 as measured in a five-day CellTiterGlo assay performed in triplicate. **c.** *In vivo* pharmacokinetic analysis of GNE-6776. Mice (3 per group) were dosed PO with 100 mg/kg or 200 mg/kg of GNE-6776. Plasma concentrations of GNE-6776 were measured at the indicated time points and plotted as a function of time. Exposure metrics relating to the free fraction EC50 for EOL-1 cells are also indicated, where target exposure = (EOL-1 IC50)/(1 - %PPB) = 1.54μM/0.066 = 23.33μM. **d.** Western blot analysis of MCF7-Ser xenografted tumors. Mice harboring MCF7-Ser xenograft tumors were treated with vehicle or 200 mg/kg GNE-6776 at 0 and 4 hours; 8 hours after the initial treatment tumors were excised and the indicated proteins were examined by immunoblotting tumor lysates. **e.** Pharmacodynamic analysis of USP7 inhibitor-treated EOL-1 xenografted tumors. Mice growing EOL-1 xenograft tumors were dosed by mouth with vehicle or 200 mg/kg GNE-6776 at 0 and 4 hours; 8 hours after the initial treatment tumors were excised and the indicated proteins were examined by immunoblotting tumor lysates. **f.** EOL-1 xenograft growth inhibition study of mice treated PO with vehicle or the indicated doses of GNE-6776. n = 7 mice per group, p-values were calculated using Dunnett's multiple comparison test. Asterisks indicate significant growth inhibition relative to vehicle-treated mice. Day 4 100 mg/kg $p = 0.0163$, Day 4 200 mg/kg $p = 0.0138$, Day 6 200 mg/kg $p = 0.0344$.

Supplementary Table 1

This table summarizes the results of the compound library combination screen performed in EOL-1 cells in the presence of either DMSO or varying concentrations of GNE-6776 or GNE-6640. Column headers denote compound used and concentration used (ie. 1000nM_6776 indicates 1000 nM of GNE-6776 treatment in combination with each compound). For each compound treatment, two summary statistics are listed that are derived from 9-point dose response curves: IC50 is the determined IC50 value of the dose response curve. This is set to the highest concentration applied if no IC50 was attained. Mean viability is the mean over all 9 dose viability measures. For each non-DMSO treatment two values are calculated to indicate synergy between the library compound and the USP7 inhibitor tested: mvdiff indicates the difference in mean viability between the DMSO treatment and the USP7 inhibitor. Positive values indicate a lower viability in the USP7-treated condition, which indicates synergism. Log2fc

indicates the log2 ratio between the IC50 value of the DMSO treatment and the USP7 treatment. Negative values indicate lower IC50s for the USP7 treatment, also signaling synergism. All unique compounds are assigned an internal gcgcid (Genentech Chemical Genomics Compound ID) for tracking purposes. Common drug names are shown in the compound_name column and either protein or functional targets of the compound, where known, are listed in the target_name column.

Methods

High-throughput screen (HTS) and counterscreen assays

USP7 UbA10 TR-FRET activity assay. Potential inhibitors of USP7 were identified in a TR-FRET-based enzyme activity assay with UbA10 as substrate. UbA10 is a fragment of the naturally occurring ubiquitin precursor, Ub52; it retains the 10 amino acid segment of Ub52 that extends beyond the ubiquitin C-terminus. The primary screening assay is a novel TR-FRET based activity assay that measures cleavage by full-length USP7 of a doubly-tagged peptide substrate. The peptide is tagged with GST on the N-terminus and with eight histidine residues on the C-terminus (GST-UbA10-His). The tags are detected by anti-GSH-d2 (TR-FRET acceptor) and anti-His-europium (TR-FRET donor), respectively. Cleavage of this substrate by USP7 at the ubiquitin C-terminus results in separation of these two tags and loss of TR-FRET signal. Compounds were dispensed into 1536-well black plates (MaKO, Aurora Microplates, Whitefish, MT) followed by 2µL full-length recombinant USP7 in assay buffer (50 mM HEPES pH 7.5, 0.1% Prionex [Pentapharm, Basel, CH], 0.01% Triton X-100, and 10 mM DTT). After a 10-minute incubation, the reaction was started by the addition of 2µL GST-UbA10-His substrate in assay buffer. After 75 minutes of reaction, 2µL of a detection antibody reagent, containing anti-His-europium (Life Technologies, Carlsbad, CA) and anti-GST-d2 (CISbio, Bedford, MA) in assay buffer were added. After 60 minutes, the fluorescence at 618 nm and 671 nm with excitation at 340 nm was read on a ViewLux reader (PerkinElmer, Waltham, MA). Cleavage of the doubly-tagged substrate resulted in loss of TR-FRET signal, while inhibition of USP7 by compound restored the signal. The TR-FRET ratio was calculated as fluorescence at 671 nm/ fluorescence at 618 nm. TR-FRET ratios were normalized to controls to determine percent inhibition for the single concentration screen. For confirmation in concentration-response mode (10 points with N = 2), percent inhibition was plotted against compound concentration, and the data were fit to a 4-parameter curve with Screener Assay Analyzer (Genedata, Basel, CH) to determine IC_{50} values. The assay buffer was optimized to maintain enzyme stability and to maximize assay signal to background: Triton X-100 was included to prevent nonspecific adsorption of the enzyme and/or substrate to the assay plate and/or to compound aggregates, Prionex carrier protein was included to help stabilize the enzyme and to prevent nonspecific adsorption to container and tubing surfaces as well as to minimize nonspecific inhibition by library compounds, and DTT served to maintain good USP7 activity and minimize the impact of inhibitors that act through redox cycling. Reagent concentrations were optimized for good assay performance at approximately 50% substrate conversion in the signal decrease assay with a key aim of minimizing the required concentration of USP7. The GST-UbA10-His concentration was adjusted to maximize assay signal, anti-GST-d2 concentration was adjusted approximately in parallel with that of the substrate, and anti-His-europium was used at a concentration that represented a minimum that is compatible with robust detection on the ViewLux plate reader. Time course evaluations were conducted to confirm that the enzyme concentration (10 nM) and reaction time (75 minutes) were in the linear range of enzyme activity. Additionally, extended time courses for the detection reaction were used to demonstrate that the 60-minute incubation was sufficient to reach equilibrium. Under the final assay conditions, reagent stability studies indicated greater than 20 hours of acceptable performance. Over the course of the screen, Z' values averaged 0.76.

USP7 di-ubiquitin FRET activity assay. Potential USP7 inhibitors were confirmed in an orthogonal activity assay with an internally quenched K63-linked di-ubiquitin substrate (U-310, Boston Biochem, Cambridge, MA). Conditions were similar to those used for the

UbA10 TR-FRET activity assay. Compounds dispensed into 1536-well plates were preincubated with full-length recombinant USP7 in assay buffer for 10 minutes, and the reaction was started by the addition of 2μL of the di-ubiquitin substrate. During the 60-minute incubation, cleavage of the substrate by USP7 resulted in the release of quenching of the TAMRA tag by the QXL tag and thus an increase in fluorescence. The fluorescence intensity was read with excitation at 540 nm and emission at 585 nm. Concentration-response assay methods and data analyses were conducted as for the UbA10 TR-FRET assay.

USP7 Ubiquitin/Rho110 activity assay. Potential USP7 inhibitors were also confirmed in an orthogonal activity assay with ubiquitin/rhodamine-110 as substrate. Compounds were preincubated for 10 minutes with 2μL full-length recombinant USP7 in assay buffer, and the reaction was started by the addition of 2μL of ubiquitin/rhodamine-110 (U-555, Boston Biochem). After a 60-minute reaction in which USP7 cleaves the rhodamine-110 from the ubiquitin and thus increases fluorescence, the fluorescence intensity was read with excitation at 485 nm and emission at 535 nm. General assay conditions and data analyses were as described above.

USP7 di-ubiquitin cleavage assay with mass spectrometric detection. To further validate compounds that met the initial HTS confirmation criteria, the ability of compounds to inhibit di-ubiquitin cleavage was assessed by monitoring the conversion to ubiquitin by mass spectrometry. Compounds were dispensed into 384-well polypropylene plates (Greiner Bio-One, Kremsmunster, AT), and 10μL USP7 were added and allowed to incubate for 10 minutes. The reaction was started by the addition of 10μL of K48-linked di-ubiquitin substrate (UC-200, Boston Biochem) and allowed to progress for 70 minutes; then it was stopped by addition of 20μL 2% formic acid. The assay plates were stored frozen at -80 °C until analysis by mass spectrometry at Agilent Technologies (Wakefield, MA). Prior to quantitation, the enzyme reaction was passed over a RapidFire cartridge to remove buffer components. Both di-ubiquitin substrate consumption and ubiquitin product formation were monitored by mass spectrometry using multiple reaction monitoring (MRM) in positive ion mode with parent ion/daughter ion transitions of 1142.1/260.1 and 770.8/817.7, respectively. Percent conversion of substrate was plotted as a function of compound concentration to generate the IC_{50} values as indicated above.

NMR screen and binding studies

All NMR spectra were recorded on Bruker Avance-600 and 800 MHz spectrometers operating at 14.1 and 18.8 Tesla using triple resonance cryogenic probes optimized for proton detection. All two-dimensional spectra were acquired with a spectral width of 16 ppm and 2048 (TROSY) or 954 (SOFAST) data points in the direct proton dimension and 28 ppm and 192 sample points in the ^{15}N dimension with echo-antiecho (TROSY) or States-TPPI (SOFAST) type selection. The resulting free induction decay resolution was 12.52 and 19.05 Hz point for the TROSY and 20.08 and 17.10 Hz/data point for the SOFAST spectra respectively. All spectra were recorded at 300 K. The pH was adjusted to 7.2 without correction due to isotope shifts. For data processing NMRPipe/NMRDraw and the BRUKER software package TOPSPIN 3.2 were used. All data evaluation was done in NMR view and CCPN. Visualization and presentation of the 3D tertiary USP7-as well as related protein structures from the RSCB Protein Database was done in Pymol (The PyMOL Molecular Graphics System, Version 1.7.4 Schrödinger, LLC). The sequential assignment of USP7 has been deposited in the Biological Magnetic Resonance Bank and can be retrieved with the accession number 26766. All samples

contained 137 mM NaCl, 10 mM Na_2HPO_4, 27 mM KCl and 1.8 mM KH_2PO_4 adjusted to a pH of 7.2 and contained 7% (w/w) ^2H_2O and 0.5 µM NaN3. Ubiquitin titration experiments were done by addition of purified bovine ubiquitin (from erythrocytes, Sigma) from a stock solution of 20mM in the same buffer. All proton chemical shifts were referenced to internal DSS (50 µM) and ^{15}N referenced indirectly using the ^1H chemical shift of the methyl group in DSS by multiplication with a factor of 0.101329118.

NMR screening conditions. The primary fragment screen was performed on 4871 fragments in mixtures of 5. The individual compound concentration was 500µM, the concentration of unlabeled USP7-CD was 7µM. Binders were identified by the presence of signals stemming from the individual ligand in the proton saturation transfer difference spectra recorded at 284K. The criterion used to identify a binder was a signal to noise ratio above 5. All primary binders were re-measured as a single compound to confirm binding. Confirmed binders were defined as compounds with an STD signal to noise ratio above 10. All confirmed binders were tested again, at 2-2.5mM, by protein observed $^1H/^{15}N$ TROSY experiments in the presence of ^{15}N USP7-CD at 220-340uM.

NMR binding studies. The concentration-dependent NMR shift perturbations caused by the interaction of unlabeled ubiquitin with labeled USP7 catalytic domain in the absence and presence of USP7 inhibitors were fit to the function for a two state fast exchanging equilibrium:

($\Delta\delta = \Delta\delta_{max} \frac{K_d+[P]+[L]-\sqrt{(K_d+[P]+[L])^2-4[P][L]}}{2[P]}$) where $\Delta\delta$ is the chemical shift change at various protein/ligand ratios, $\Delta\delta_{max}$ is the chemical shift change at saturation, K_d is the dissociation constant, and [L] and [P] are the ligand and protein concentrations, respectively. The chemical shift change is the root mean square of the ^1H and ^{15}N values scaled by 1 and 0.15 respectively. Seven ubiquitin concentrations were measured (0, 100, 250, 500, 750, 1000, 2500uM). USP7 inhibitors were added at 1mM. The mean values and standard deviation were calculated by averaging the values obtained for eight well-resolved cross peaks: Y339, S341, D342, G375, A381, G382, D412, and I419. Data were visualized using GraphPad Prism.

Compound synthesis

Synthesis of GNE-6640 is representative of the syntheses for GNE-6776, GNE-6641, GNE-2931, GNE-2917, and GNE-2916 .

3,5-Dibromo-4-ethylpyridin-2-amine: Into a 500-mL 3-necked round-bottom flask purged and maintained with an inert atmosphere of nitrogen was placed 4-ethylpyridin-2-amine (10 g, 81.85 mM), tetrahydrofuran (200 mL), and NBS (29 g, 162.94 mmol) at 0 ^oC. The resulting solution was stirred at room temperature for 15 min and then concentrated under vacuum. The residue was purified on a silica gel column eluting with DCM/MeOH (100:1-20:1) to afford 18 g (79%) of the title compound. ^1H NMR (400 MHz, $CDCl_3$) δ 8.04 (s, 1H), 7.33 – 7.23 (m, 0H), 4.93 (s, 2H), 2.93 (q, *J* = 7.5 Hz, 2H), 1.60 (s, 0H), 1.17 (t, *J* = 7.5 Hz, 3H). LCMS (ESI M/Z): 264.1 ($M + H^+$).

3-Bromo-4-ethylpyridin-2-amine: Into a 500-mL 3-necked round-bottom flask purged and maintained with an inert atmosphere of nitrogen was placed 3,5-dibromo-4-ethylpyridin-2-amine (18 g, 64.29 mmol) in tetrahydrofuran (300 mL). To this was added a solution of n-BuLi (in hexane) (58 mL, 2.2 mol/L) at -78 ^oC. The resulting solution was stirred at -78^oC for 1 h, quenched by the addition of 450 mL NH_4Cl and then extracted with ethyl acetate (2 x 500 mL). The combined organic layers were washed with brine (2 x 500 mL), dried over anhydrous sodium sulfate and concentrated under vacuum. The crude product was purified by Flash-Prep-HPLC to afford 12 g (93%) of the title

compound as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.89 (d, *J* = 5.1 Hz, 1H), 6.60 – 6.48 (m, 1H), 4.93 (s, 2H), 2.69 (q, *J* = 7.6 Hz, 2H), 1.33 – 1.14 (m, 5H). LCMS (ESI M/Z): 264.1 (M + H^+).

4-Ethyl-3-(4methoxyphenyl)pyridn-2-amine: Into a 500-mL 3-necked round-bottom flask purged and maintained with an inert atmosphere of nitrogen was placed a solution of 3-bromo-4-ethylpyridin-2-amine (12 g, 59.68 mmol) in CH_3CN (100 mL), (4-methoxyphenyl)boronic acid (11 g, 72.39 mmol), Na_2CO_3(120 mL, sat.), and Pd(dppf)Cl_2 (1.2 g, 1.64 mmol). The resulting solution was stirred at 110 ^{o}C for 1 h, diluted with of 500 mL of EA and then extracted with of ethyl acetate (2 x 500 mL). The combined organic layers were washed with brine (3 x 200 mL), dried over anhydrous sodium sulfate and concentrated under vacuum. The residue was purified on a silica gel column eluting with ethyl acetate/petroleum ether (1:100-1:10) to afford 10 g (73%) of the title compound. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.06(d, *J* = 5.1 Hz, 1H), 7.21 – 7.11 (m, 2H), 7.06 – 6.97 (m, 2H), 6.64 (d, *J* = 5.4 Hz, 1H), 4.48 (s, 2H), 3.86 (s, 3H), 2.31 (q, *J* = 7.6 Hz, 2H), 1.11 – 0.88 (m, 3H). LCMS (ESI M/Z): 264.1 (M + H^+).

5-Bromo-4-ethyl-3-(4-methoxyphenyl)pyridin-2-amine: Into a 250-mL 3-necked round-bottom flask purged and maintained with an inert atmosphere of nitrogen was placed 4-ethyl-3-(4methoxyphenyl)pyridn-2-amine (10 g, 43.80 mmol), THF (100 mL), followed by NBS (7.8 g, 43.83 mmol) at 0 ^{o}C. The resulting solution was stirred at room temperature for 15 min, diluted with 500 mL EtOAc and 500 mL H_2O. The resulting solution was extracted with ethyl acetate (2 x 500 mL). The organic layers were combined, washed with brine (2 x 500 mL) and concentrated under vacuum. The residue was purified on a silica gel column eluting with ethyl acetate/petroleum ether (1:20-1:10) to afford 8 g (59%) of the title compound. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.09 (s, 1H), 7.22 – 7.09 (m, 2H), 7.09 – 6.98 (m, 2H), 4.55 (s, 2H), 2.47 (q, *J* = 7.5 Hz, 2H), 1.10 – 0.84 (m, 3H). LCMS (ESI M/Z): 264.1 (M + H^+).

4-(2-Amino-5-bromo-4-ethylpyridin-3-yl)phenol: Into a 250-mL 3-necked round-bottom flask purged and maintained with an inert atmosphere of nitrogen was placed 5-bromo-4-ethyl-3-(4-methoxyphenyl)pyridin-2-amine as a white solid (8 g, 26.04 mmol), dichloromethane (100 mL), followed by tribromoborane (19.6 g, 78.24 mmol) at 0 ^{o}C. The resulting solution was stirred at room temperature for 1 h and then quenched by the addition of 100 mL of $NaHCO_3$ (1 M) at 0 ^{o}C. The solids were collected by filtration and then washed with 100 mL H_2O and 300 mL of EA/PE (1:1) to afford 6.3 g (83%) of the title compound as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.09 (s, 1H), 7.22 – 7.09 (m, 2H), 7.09 – 6.98 (m, 2H), 4.55 (s, 2H), 2.47 (q, *J* = 7.5 Hz, 2H), 1.10 – 0.84 (m, 3H). LCMS (ESI M/Z): 264.1 (M + H^+).

4-[2-Amino-4-ethyl-5-(1*H*-indazol-5-yl)-3-pyridyl]phenol (GNE-6640): Into a 250-mL 3-necked round-bottom flask purged and maintained with an inert atmosphere of nitrogen was placed 4-(2-amino-5-bromo-4-ethylpyridin-3-yl)phenol (1.0 g, 3.41 mmol,), 6-(tetramethyl-1,3,2-dioxaborolan-2-yl)-2*H*-indazole (880 mg, 3.41 mmol,), potassium carbonate (3.3 g, 23.88 mmol), water (30 mL), 1,4-dioxane (25 mL), and Pd(dppf)Cl_2 (200 mg, 0.3 mmol). The resulting solution was stirred at 80 ^{o}C for 16 h, diluted with 500 mL H_2O and 500 mL ethyl acetate. The organic layer was washed 30 with brine (2 x 250 mL) and concentrated under vacuum. The residue was purified on a silica gel column eluting with DCM/CH_3OH (20:1-10:1) to afford the titled compound. ^{1}H NMR (400 MHz, DMSO-d_6) δ 13.07 (s, 1H), 9.52 (s, 1H), 8.07 (d, *J* = 1.0 Hz, 1H), 7.74 (s, 1H), 7.65 (dd, *J* = 1.6, 0.8 Hz, 1H), 7.56 (dt, *J* = 8.6, 0.9 Hz, 1H), 7.28 (dd, *J* = 8.5, 1.6 Hz, 1H), 7.10 – 7.04 (m, 2H), 6.92 – 6.86 (m, 2H), 4.94 (s, 2H), 2.26 (q, *J* = 7.4 Hz, 2H), 0.60 (t, *J* = 7.4 Hz, 3H). LCMS (ESI M/Z): 331.1 (M+H). HRMS m/e 331.1533. (M + H^+, $C_{20}H_{19}ON_4$ requires 331.1653)

5-[6-Amino-4-ethyl-5-(4-hydroxyphenyl)-3-pyridyl]-*N*-methyl-pyridine-2-carboxamide (GNE-6776): ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.56 (s, 1H), 8.80 (q, *J* = 4.7 Hz, 1H), 8.59 (dd, *J* = 2.3, 0.9 Hz, 1H), 8.07 (dd, *J* = 8.0, 0.9 Hz, 1H), 7.96 (dd, *J* = 8.0, 2.2 Hz, 1H), 7.79 (s, 1H), 7.11 – 7.02 (m, 2H), 6.94 – 6.86 (m, 2H), 5.19 (s, 2H), 2.84 (d, *J* = 4.8 Hz, 3H), 2.26 (q, *J* = 7.4 Hz, 2H), 0.62 (t, *J* = 7.4 Hz, 3H). HRMS m/e 349.1659. (M + H^+, $C_{20}H_{21}O_2N_4$ requires 349.1650)

4-[2-Amino-4-ethyl-5-(2-methylindazol-6-yl)-3-pyridyl]phenol (GNE-6641). ^{1}HNMR (400MHz, DMSO-d_6) δ 9.54 (s, 1H), 7.81 (dd, *J* = 7.3, 1.5Hz, 2H), 7.77 – 7.52 (m, 4H), 7.13 – 7.01 (m, 2H), 6.93 – 6.85 (m, 2H), 5.10 (s, 2H), 3.28 (s, 2H), 2.24 (q, *J* = 7.4Hz, 2H), 0.60 (t, *J* = 7.5Hz, 3H). HRMS m/e 345.1710. (M + H^+, $C_{21}H_{21}ON_4$ = 345.1650)

4-[2-Amino-4-ethyl-5-(2-methoxyphenyl)-3-pyridyl]phenol (GNE-2931). ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.54 (s, 1H), 8.00 (s, 2H), 7.90 – 7.77 (m, 3H), 7.73 (s, 1H), 7.54 – 7.40 (m, 2H), 7.39 – 7.34 (m, 1H), 7.12 – 7.00 (m, 2H), 6.94 – 6.82 (m, 2H), 3.03 (s, 3H), 2.91 (q, *J* = 4.4 Hz, 2H), 1.71 (t, *J* = 4.4 Hz, 3H). HRMS m/e 321.1598 (M + H^+, $C_{20}H_{21}O_2N_2$ requires 321.1650)

3-[6-Amino-4-ethyl-5-(4-hydroxyphenyl)-3-pyridyl]benzamide (GNE-2917). ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.52 (s, 1H), 8.48 (s, 2H), 7.58 (s, 1H), 7.34 (ddd, *J* = 8.2, 7.3, 1.8 Hz, 1H), 7.13 (dd, *J* = 7.4, 1.8 Hz, 1H), 7.10 – 6.94 (m, 6H), 6.88 (d, J = 7.6 Hz, 2H), 2.91 (q, *J* = 4.4 Hz, 2H), 1.71 (t, *J* = 4.4 Hz, 3H). HRMS m/e 334.1550 (M + H^+, $C_{20}H_{20}O_2N_3$ requires 334.1700)

4-[6-Amino-4-ethyl-5-(4-hydroxyphenyl)-3-pyridyl]benzamide (GNE-2916). ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.53 (s, 1H), 7.97 (s, 1H), 7.95 – 7.87 (m, 1H), 7.72 (s, 3H), 7.43 – 7.35 (m, 1H), 7.33 (s, 1H), 7.12 – 7.00 (m, 2H), 6.93 – 6.83 (m, 2H), 5.04 (s, 2H), 2.27 (q, *J* = 7.4 Hz, 2H), 0.61 (t, *J* = 7.5 Hz, 3H). LCMS (ESI) m/z 334.2 [M+H^+]

Synthesis of GNE-6831 is representative of the syntheses for GNE-2090, GNE-2143, and GNE-2148.

Synthesis of 9-chloro-*N*-(5-chloro-2,4-dimethoxy-phenyl)-*N*-(cyanomethyl)-5,6,7,8-tetrahydroacridine-3-carboxamide (GNE-6831). 9-chloro-5,6,7,8-tetrahydroacridine-3-carboxylic acid (100 mg, 0.38 mmol) was dissolved in 2 mL of DMF and charged with HATU (144 mg, 0.38 mmol). After stirring at RT for minutes, the mixture was then charged with 3-((5-chloro-2,4-dimethoxyphenyl)amino)propanenitrile (91 mg, 0.38 mmol). The resulting solution was stirred at 110 ^{o}C for 1 h, diluted with of 50 mL of EA and then extracted with ethyl acetate (2 x 100 mL). The combined organic layers were washed with brine (3 x 200 mL), dried over anhydrous sodium sulfate and concentrated under vacuum. The residue was purified on a silica gel column eluting with ethyl acetate/petroleum ether (1:100-1:10) to afford the title compound (110 mg, 0.30 mmol, 78% yield) of the title compound. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.02 (d, *J* = 8.7 Hz, 1H), 7.82 (d, *J* = 1.7 Hz, 1H), 7.65 – 7.56 (m, 2H), 6.67 (s, 1H), 4.84 (s, 2H), 3.79 (s, 3H), 3.73 (s, 3H), 2.92 (m, 4H), 1.85 (m, 4H). HRMS m/e 470.1033 (M + H^+, $C_{24}H_{22}O_3N_3Cl_2$, requires 470.1029)

9-Chloro-*N*-(5-chloro-2,4-dimethoxy-phenyl)-*N*-(2,3-dihydroxypropyl)-5,6,7,8-tetrahydroacridine-3-carboxamide (GNE-2090): ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.96 (dd, *J* = 8.6, 3.4 Hz, 2H), 7.77 – 7.68 (m, 2H), 7.66 – 7.47 (m, 1H), 6.56 (d, *J* = 14.2 Hz, 1H), 4.99 (d, *J* = 5.2 Hz, 1H), 4.54 (tt, *J* = 15.4, 5.8 Hz, 2H), 3.46 – 3.33 (m, 4H), 2.91 (m, 4H), 1.85 (m, 4H). LCMS (ESI) m/z 505.2 [M + H^+].

9-Chloro-*N*-(2-hydroxyethyl)-*N*-methyl-5,6,7,8-tetrahydroacridine-3-carboxamide (GNE-2143): ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.16 (d, *J* = 8.2 Hz, 1H), 7.95 (d, *J* = 8.2 Hz, 1H), 7.7 (m, 1H), (3.66 (s, 3H), 3.57 (s, 1H), 1.90 (td, *J* = 3.8, 1.8 Hz, 4H). HRMS m/e 319.1208(M + H^+, $C_{17}H_{20}O_2N_2Cl$ requires 319.1500)

9-Chloro-*N*-methyl-*N*-(2-morpholino-2-oxo-ethyl)-5,6,7,8-tetrahydroacridine-3-carboxamide (GNE-2148): ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.21 (d, *J* = 8.6 Hz, 1H),

8.16 (d, *J* = 8.6 Hz, 1H), 7.96 – 7.90 (m, 1H), 4.40 (s, 1H), 3.62 (s, 4H), 3.54 – 3.41 (m, 4H), 3.15 (s, 3H), 3.00 (t, 20.9 Hz, 4H), 1.94 – 1.86 (m, 4H). HRMS m/e 402.1579 (M + H^+, $C_{21}H_{25}O_3N_3Cl$ requires 402.1725)

Synthesis of (1*S*)-1-(5-bromo-1*H*-indol-2-yl)-*N*-(2-phenylethyl)ethanamine (GNE-0300). 1-(5-bromo-1*H*-indol-2-yl)ethan-1-one (90 mg, 0.38 mmol) was dissolved in 2 mL of DMF and charged with (*S*)-phenethylamine (46 mg, 0.38 mmol). After stirring at RT for 30 minutes, the mixture was diluted with 50 mL ethyl acetate and 50 mL water. The mixture was partitioned and the organic was collected. The aqueous was then extracted with of ethyl acetate (2 x 100 mL). The combined organic layers were washed with brine (3 x 200 mL), dried over anhydrous sodium sulfate and concentrated under vacuum. The residue was purified on a silica gel column eluting with ethyl acetate/petroleum ether (1:100-1:10) to afford the title compound (120 mg, 0.35 mmol, 92% yield) of the titled compound. ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.05 (s, 1H), 7.59 (d, *J* = 1.9 Hz, 1H), 7.30 – 7.19 (m, 3H), 7.20 – 7.06 (m, 4H), 6.21 (d, *J* = 1.8 Hz, 2H), 3.93 (q, *J* = 6.6 Hz, 1H), 2.77 – 2.56 (m, 4H), 1.36 (d, *J* = 6.6 Hz, 3H). LCMS (ESI) m/z 343.9 [M + H^+]

Synthesis of GDC-0570 follows similar procedures as described for GDC-0339 as below.
(*Z*)-2,3,6,7-Tetrahydro-1*H*-azepine hydrochloride: 4N Hydrogen chloride in 1,4-dioxane (250 mL; 1 mol) was added over 5 minutes to a stirred, ice cooled solution of (*Z*)-*tert*-butyl 2,3,6,7-tetrahydro-1*H*-azepine-1-carboxylate (50 g; 0.254 mol) in methanol (250 mL). On complete addition, the ice bath was removed and stirring continued at room temperature for 3.75 h. Volatiles were removed under reduced pressure and the residue triturated twice with diethyl ether (300 mL) to afford (*Z*)-2,3,6,7-tetrahydro-1*H*-azepine hydrochloride as a pale pink solid (32.3 g; 95%). ^{1}H-NMR (DMSO-d_6, 400 MHz) δ 9.54 (br s, 2H), 6.40-6.25 (m, 2H), 3.15-3.05 (m, 4H), 2.55-2.40 (m, 4H).

(*Z*)-1-(1-Methyl-4-nitro-1*H*-pyrazol-5-yl)-2,3,6,7-tetrahydro-1*H*-azepine: A mixture of (*Z*)-2,3,6,7-tetrahydro-1*H*-azepine hydrochloride (32.3 g; 0.24 mol), 5-chloro-1-methyl-4-nitro-1*H*-pyrazole (37.2 g; 0.23 mol), potassium fluoride (56.24 g; 0.96 mol) and diisopropylethylamine (64 mL; 0.362 mol) in anhydrous DMSO (650 mL) was heated at 75 °C for 21 h. On cooling, the mixture was poured into water (1500 mL), extracted with ethyl acetate (4 x 500 mL) and the combined organics washed with water (2 x 400 mL), brine (300 mL) and dried ($MgSO_4$). The solvent was removed under reduced pressure to afford (*Z*)-1-(1-methyl-4-nitro-1*H*-pyrazol-5-yl)-2,3,6,7-tetrahydro-1*H*-azepine as a light brown solid (50.74 g; 99%). ^{1}H-NMR ($CDCl_3$, 400 MHz) δ 8.00 (s, 1H), 5.95-5.85 (m, 2H), 3.80 (s, 3H), 3.30-3.20 (m, 4H), 2.45-2.35 (m, 4H). LCMS (ESI M/Z): 223.1 [M + H^+].

4-(1-Methyl-4-nitro-1*H*-pyrazol-5-yl)-8-oxa-4-azabicyclo[5.1.0]octane: 77% meta-Chloroperbenzoic acid (77 g; 0.343 mol) was added portion-wise over 10 minutes to a stirred, ice cooled solution of (*Z*)-1-(1-methyl-4-nitro-1*H*-pyrazol-5-yl)-2,3,6,7-tetrahydro-1*H*-azepine (50.74 g; 0.23 mol) in dichloromethane (1000 mL). Ice bath used to control minor exotherm observed during a smaller scale reaction. On complete addition, the ice bath was removed and stirring continued at room temperature for 18 h. The reaction mixture was washed with saturated sodium hydrogen carbonate (750 mL), 1N sodium hydroxide (2 x 500 mL) and brine (350 mL). The organics were dried ($MgSO_4$) and the solvent removed under reduced pressure to afford 4-(1-methyl-4-nitro-1*H*-pyrazol-5-yl)-8-oxa-4-azabicyclo[5.1.0]octane as a pale yellow solid (55.6 g). ^{1}H-NMR ($CDCl_3$, 400 MHz) δ 7.98 (s, 1H), 3.75 (s, 3H), 3.50-3.35 (m, 2H), 3.30-3.20 (m, 2H), 2.95-2.80 (m, 2H), 2.35-2.15 (m, 4H). LCMS (ESI M/Z): 239.2 [M + H^+].

rel-(4*R*,5*R*)-5-Azido-1-(1-methyl-4-nitro-1*H*-pyrazol-5-yl)azepan-4-ol: To a stirred solution of 4-(1-methyl-4-nitro-1*H*-pyrazol-5-yl)-8-oxa-4-azabicyclo[5.1.0]octane (29.28 g) in methanol (350 mL) and water (90 mL) was added ammonium chloride (16.5 g; 0.308 mol) followed by sodium azide (20 g; 0.307 mol). The mixture was heated behind a blast screen at 70 °C for 22 h, cooled then concentrated to 100 mL under reduced pressure at 40 °C. The concentrated solution was poured into water (1300 mL), extracted with dichloromethane (4 x 400 mL) and the combined organics dried ($MgSO_4$). Evaporation under reduced pressure at 35 °C gave rel-(4*R*,5*R*)-5-azido-1-(1-methyl-4-nitro-1*H*-pyrazol-5-yl)azepan-4-ol (anti-isomer and a racemic mixture) as a pale yellow oil (34.5 g; 96% over 2 steps). ^{1}H-NMR ($CDCl_3$, 400 MHz) δ 8.03 (s, 1H), 3.85-3.77 (m, 1H), 3.77 (s, 3H), 3.65-3.55 (m, 1H), 3.45-3.15 (m, 4H), 2.85-2.70 (m, 1H), 2.25-2.10 (m, 2H), 2.05-1.85 (m, 2H). LCMS (ESI M/Z): 282.1 [M + H^+].

rel-(4*R*,5*R*)-4-Azido-5-fluoro-1-(1-methyl-4-nitro-1*H*-pyrazol-5-yl)azepane: 50% Deoxofluor in THF (111 mL; 0.307 mol) was added slowly over 20 minutes to a stirred, ice cooled solution of rel-(4*R*,5*R*)-5-azido-1-(1-methyl-4-nitro-1*H*-pyrazol-5-yl)azepan-4-ol (32.5 g; 0.115 mol) in dichloromethane (500 mL). On complete addition, the ice bath was removed and stirring continued at room temperature for 20 h. The reaction mixture was re-cooled in an ice bath and saturated sodium hydrogen carbonate (400 mL) added dropwise (effervescence!). After stirring for 30 minutes the layers were separated and the aqueous layer extracted with dichloromethane (2 x 500 mL). Pooled organics were dried ($MgSO_4$) and the solvent removed under reduced pressure. Flash column chromatography on silica eluting with 0 – 100% ethyl acetate in isohexane gradient afforded rel-(4*R,*5*R*)-4-azido-5-fluoro-1-(1-methyl-4-nitro-1*H*-pyrazol-5-yl)azepane as a pale orange oil (26 g; 80%). ^{1}H-NMR ($CDCl_3$, 400 MHz) δ 8.03 (s, 1H), 4.90-4.65 (m, 1H), 4.00-3.85 (m, 1H), 3.77 (s, 3H), 3.40-3.10 (m, 4H), 2.35-2.05 (m, 3H), 1.95-1.75 (m, 1H). LCMS (ESI M/Z): 284.3 [M + H^+].

rel-(4*R*,5*R*)-5-Fluoro-1-(1-methyl-4-nitro-1*H*-pyrazol-5-yl)azepan-4-amine: A mixture of rel-(4*R*,5*R*)-4-azido-5-fluoro-1-(1-methyl-4-nitro-1*H*-pyrazol-5-yl)azepane (26 g; 91.8 mmol) and triphenylphosphine (24.1 g; 92 mmol) in tetrahydrofuran (400 mL) and water (80 mL) was heated at 60 °C for 20 h, cooled and concentrated to approximate 80 mL under reduced pressure. Ethyl acetate (500 mL) was added and the mixture extracted with 1N HCl (4 x 125 mL). Pooled acidic extracts were washed with ethyl acetate (500 mL), basified to pH 14 with 6N NaOH and extracted with dichloromethane (3 x 400 mL). Combined extracts were dried ($MgSO_4$) and the solvent removed under reduced pressure to give rel-(4*R*,5*R*)-5-fluoro-1-(1-methyl-4-nitro-1*H*-pyrazol-5-yl)azepan-4-amine as a pale yellow oil (22 g; 93%). ^{1}H-NMR ($CDCl_3$, 400 MHz) δ 8.03 (s, 1H), 4.60-4.40 (m, 1H), 3.77 (s, 3H), 3.45-3.10 (m, 5H), 2.35-1.90 (m, 3H), 1.80-1.65 (m, 1H), 1.60 (br s, 2H). LCMS (ESI M/Z): 258.3 [M + H^+].

***tert*-Butyl (4*R*,5*R*)-5-fluoro-1-(1-methyl-4-nitro-1*H*-pyrazol-5-yl)azepan-4-ylcarbamate:** Di-*tert*-butyl dicarbonate (28 g; 128.3 mmol) was added to a stirred, ice cooled solution of rel-(4*R*,5*R*)-5-fluoro-1-(1-methyl-4-nitro-1*H*-pyrazol-5-yl)azepan-4-amine (22 g; 85.6 mmol) and diisopropylethylamine (22.4 mL; 128.6 mmol) in dichloromethane (600 mL). On complete addition, the ice bath was removed and stirring continued at room temperature for 20 h. The reaction mixture was washed with saturated sodium hydrogen carbonate (500 mL) and the aqueous layer re-extracted with dichloromethane (2 x 300 mL). Pooled organics were dried ($MgSO_4$) and the solvent removed under reduced pressure. Flash column chromatography on silica eluting with 0

– 100% ethyl acetate in isohexane gradient afforded a pale yellow solid (29.7 g; 97%). Chiral separation of the racemic mixture by supercritical fluid chromatography (SFC) using Chiralpak IA column with an isocratic mobile phase of 15% methanol (with 0.1% NH_4OH) in carbon dioxide gave the desired product (second peak) as a single enantiomer. ^{1}H-NMR ($CDCl_3$, 400 MHz) δ 8.04 (s, 1H), 5.05 (m, 1H), 4.80-4.55 (m, 1H), 4.15-4.05 (m, 1H), 3.79 (s, 3H), 3.45-3.10 (m, 4H), 2.35-2.05 (m, 3H), 1.95-1.80 (m, 1H), 1.47 (s, 9H). LCMS (ESI M/Z): 358.3 [M + H^+].

***tert*-Butyl (4*R*,5*R*)-5-fluoro-1-(4-amino-1-methyl-1*H*-pyrazol-5-yl)azepan-4-ylcarbamate:** To a stirred solution of *tert*-butyl (4*R*,5*R*)-5-fluoro-1-(1-methyl-4-nitro-1*H*-pyrazol-5-yl)azepan-4-ylcarbamate (15.0 g; 42 mmol) in ethanol (1000 mL) and water (100 mL) was added ammonium chloride (11.34 g; 210 mmol) and iron powder (9.4 g; 168 mmol). The mixture was heated at 99 °C for 2.75 h, cooled, filtered through Celite® and evaporated to approximately 100 mL. The concentrate was diluted with water (1000 mL) and extracted with ethyl acetate (2 x 500 mL). The pooled extracts were washed with water (200 mL), dried ($MgSO_4$) and the solvent removed under reduced pressure to give *tert*-butyl (4*R*,5*R*)-5-fluoro-1-(4-amino-1-methyl-1*H*-pyrazol-5-yl)azepan-4-ylcarbamate as a pale brown solid (12.6 g; 91%). ^{1}H-NMR ($CDCl_3$, 400MHz) δ 7.13 (s, 1H), 6.35-6.20 (m, 1H), 4.85-4.65 (m, 1H), 4.35-4.15 (m, 1H), 3.66 (s, 3H), 3.45-3.30 (m, 2H), 3.15-2.90 (m, 2H), 2.65 (s, 2H), 2.30-2.15 (m, 1H), 2.15-1.95 (m, 2H), 1.90-1.80 (m, 1H), 1.45 (s, 9H). LCMS (ESI M/Z): 327.2 [M + H^+].

***tert*-Butyl (4*R*,5*R*)-1-(4-(5-*tert*-butoxycarbonylamino-2-(2,6-difluorophenyl)thiazole-4-carboxamido)-1-methyl-1*H*-pyrazol-5-yl)-5-fluoroazepan-4-ylcarbamate:** A mixture of *tert*-butyl (4*R*,5*R*)-5-fluoro-1-(4-amino-1-methyl-1*H*-pyrazol-5-yl)azepan-4-ylcarbamate (12.6 g, 38.5 mmol), 5-(*tert*-butoxycarbonylamino)-2-(2,6-difluorophenyl)thiazole-4-carboxylic acid (14.4 g, 40.4 mmol), diisopropylethylamine (13.4 mL, 77 mmol) and PyBOP (26.1 g, 50 mmol) in dichloromethane (400 mL) was stirred at room temperature for 48 h. Saturated sodium hydrogen carbonate (600 mL) was added and stirring continued for 0.5 h. The mixture was filtered, the layers separated and the aqueous extracted with dichloromethane (500 mL). The pooled organics were dried ($MgSO_4$) and the solvent removed under reduced pressure. Flash column chromatography on silica eluting with 0 – 100% ethyl acetate in isohexane gradient afforded *tert*-butyl (4*R*,5*R*)-1-(4-(5-*tert*-butoxycarbonylamino-2-(2,6-difluorophenyl)thiazole-4-carboxamido)-1-methyl-1H-pyrazol-5-yl)-5-fluoroazepan-4-ylcarbamate as a pale yellow solid (2.5 g, 9.8%). Further elution with 0-10% methanol in ethyl acetate, then 10% methanol in dichloromethane, evaporation of relevant fractions under reduced pressure and trituration of the residue with cold diethyl ether gave a further 7.61g (30%) of product. ^{1}H-NMR ($CDCl_3$, 400MHz) δ 10.34 (s, 1H), 8.75 (s, 1H), 7.89 (s, 1H), 7.40-7.25 (m, 1H), 7.15-7.00 (m, 2H), 4.95-4.85 (m, 1H),4.85-4.65 (m, 1H), 4.15-4.00 (m, 1H), 3.77 (s, 3H), 3.45-3.30 (m, 2H), 3.25-3.00 (m, 2H), 2.35-2.10 (m, 3H), 1.95-1.75 (m, 1H), 1.55 (s, 9H), 1.43 (s, 9H).

5-Amino-*N*-(5-((4*R*,5*R*)-4-amino-5-fluoroazepan-1-yl)-1-methyl-1*H*-pyrazol-4-yl)-2-(2,6-difluorophenyl)thiazole-4-carboxamide (GDC-0339): *tert*-butyl (4*R*,5*R*)-1-(4-(5-*tert*-butoxycarbonylamino-2-(2,6-difluorophenyl)thiazole-4-carboxamido)-1-methyl-1*H*-pyrazol-5-yl)-5-fluoroazepan-4-ylcarbamate (10.11 g, 15.2 mmol) in 4N HCl in dioxane (200 mL) and methanol (200 mL) was stirred at room temperature for 20 h. Evaporation under reduced pressure afforded a pale brown solid which was dissolved in 50% methanol in dichloromethane and added to an SCX cartridge (strong cation exchange chromatography). After washing with dichloromethane and methanol, elution with 1N

ammonia in methanol and evaporation of the eluent under reduced pressure afforded 5-amino-*N*-(5-((4*R*,5*R*)-4-amino-5-fluoroazepan-1-yl)-1-methyl-1*H*-pyrazol-4-yl)-2-(2,6-difluorophenyl)thiazole-4-carboxamide as the free base (6.3 g, 89%). ^{1}H-NMR (DMSO-d_6, 400MHz) δ 8.92 (s, 1H), 7.65-7.50 (m, 4H), 7.40-7.25 (m, 2H), 4.60-4.45 (m, 1H), 3.70 (s, 3H), 3.35-3.05 (m, 5H), 2.30-1.60 (m, 6H). LCMS (ESI M/Z): 466.1 [M + H$^+$].

Hit-to-lead selection cascade assays

Biochemical deubiquitinase assays. Biochemical deubiquitinase assays used Ubiquitin-Rho110 as a substrate to enable kinetic monitoring of reactions for USP7, USP7 catalytic domain, USP5, and USP47. The deubiquitinase proteins and their concentrations that were used in biochemical reactions were as follows: USP7, full-length N-terminal His-tag, native C-Terminus: 0.18 nM (Genentech, Hs_USP7 2-1102); USP7 catalytic domain N-terminal His-tag: 40 nM (Genentech, Hs_USP7.K208-K554); USP5; full-length: 0.5 nM (Boston Biochem cat # E-322, lot 02010210); USP47, full-length N-terminal His-tag:0.8 nM (Genentech Hs_USP47.M1-D1287). Substrate Km values for USP proteins were: USP7, full-length, 2 μM; USP5, full-length, 0.33 μM; USP47 full-length, 0.175 μM. The final assay conditions were as follows: The Reaction Buffer consisted of 50 mM Tris (pH 7.5), 0.01%(v/v) Triton X-100, 2.5 mM Dithiothreitol, 0.1% (w/v) bovine gamma globulin (Sigma cat # G5009-25G). The final substrate Ubiquitin-Rho110 (Boston Biochem cat # U-555) concentration used for reactions was 1μM. Reactions were carried out for 1 hour at room temperature, in black 20μL volume polystyrene ProxiPlate 384 F Plus (PerkinElmer cat # 6008260). Test compounds, including a control USP7 inhibitor (Ub-aldehyde, Boston Biochem cat # U-201) were serially diluted in DMSO, in 384-well clear V-bottom polypropylene plates (Greiner cat # 781280). Compounds in DMSO were diluted 10-fold into Reaction Buffer, to achieve 3-fold the final desired concentration. The substrate, Ubiquitin-Rho110 (Boston Biochem cat # U-555), was prepared at 3μM (3-fold the final concentration) and 5μL was dispensed into the reaction plate. Five μL of the compounds (diluted in Reaction Buffer at 3-fold the final concentration) were transferred to the reaction plate. Five μL of DUB protein, which was diluted in Reaction Buffer at 3-fold the final concentration, was transferred to the reaction plate to initiate the reaction. After 1 hour incubation at room temperature the reaction was quenched by the addition of 5μL of 400mM acetic acid, in the case of an endpoint reaction. The enzymatic product was measured by quantifying the fluorescence signal of cleaved Rhodamine-110 using excitation at 485 nm and emission at 535nm. When pre-incubation of DUB with compounds was required, the order of addition of reagents was modified to pre-mix the compounds with DUB (with a 1 hour incubation period), prior to the addition of the substrate and the initiation of the reaction period. Percentage inhibition values were calculated relative to a no enzyme control and an uninhibited enzyme control. Curve fitting and IC_{50} calculations were carried out using Genedata Screener software. Conversion between IC_{50} and Ki values were carried out using the Cheng-Prusoff equation.

Cellular ubiquitin-MDM2 assay. HCT116 colon cancer cells or SJSA-1 osteosarcoma cells were seeded at a density of 150,000 cell per well in 90μl (RPMI 1640 media, 10% FBS (or 0.5% FBS for low serum conditions), 1X GlutaMAX™ from Gibco) in 96-well black clear bottom, TC-treated (Greiner, Cat# 655090), and incubated for 2 hours at 37°C, 5% CO_2 in a tissue culture incubator. Compounds were prepared in a serial dilution in DMSO at 200x the final desired concentration in a 96-well polypropylene V-bottom (Greiner, Cat# 651261), then diluted 1:20 in RPMI tissue culture medium and 10μl transferred to each well of the cell plate. Cell plates were incubated overnight for 20

hours, 37°C, 5% CO_2. Twenty μl of a 120 μM stock (in RPMI) of the proteasome inhibitor, MG132 (Cayman Chemical, Cat# 10012628), was added to each well. Cells were incubated for 1 hour at 37°C, 5% CO_2. Quantitation of ubiquitin-MDM2 was carried out using Ub/Total MDM2 whole cell lysate kit (MSD, Cat# K15168D-2). Cells were lysed by adding 15μl of 5x MSD lysis buffer (containing additives: 10mM NaF, 10mM beta-glycerophosphate, 1.5mM Na_3VO_4, protease inhibitor cocktail (Sigma, P8340) to each well and incubated at 4°C for 30 minutes with shaking. One hundred μl of lysate was transferred to each well of the MSD 96-well plate, incubated at room temperature for 1 hour while shaking (650 RPM) in the dark. The MSD plates were washed 3 times in Tris buffered saline (50 mM Tris-Cl, pH 7.5. 150 mM NaCl) using a Biotek EL405 plate washer. Three mL of detection antibody solution was prepared per plate (1 mL of block buffer A, 1.82ml 1X Tris wash buffer, 150μl 2% Blocker D-M, 30μl 10% Blocker D-R, 60μl 50X anti-total MDM2 antibody). Twenty-five μl of detection antibody solution was added per well and incubated for 1 hour at room temperature (650RPM) in the dark. Plates were washed 3 times in Tris buffered saline using a Biotek EL405 plate washer. MSD read buffer was prepared according to manufacturer's instructions and 150 μl added per well. Plates were read using a MSD Sector Reader. The final measurement was the ratio of ubiquitinated MDM2 / total MDM2. Percentage increase in ubiquitin-MDM2 was calculated relative to DMSO controls using Genedata Screener software.

Total MDM2 immunofluorescence. HCT-116 cells were seeded at a density of 40,000/well in 50μL/well in 384 well tissue culture plates (Greiner #781091) in RPMI, 10% FBS, 2mM L-glutamine, and incubated overnight. Test compounds were prepared in a 20-point serial dilution (1:2-fold) in DMSO using a Biomek FX in a 384 well Labcyte-approved polypropylene plate (Labcyte P05525). Compounds were acoustically dispensed into the cell plates using a Labcyte Echo (final total volume transferred was 50nL (DMSO final concentration was 0.1%v/v). Cell plates were incubated at 37°C 5% CO_2, for 24 hours. Cells were fixed by addition of 15μL of 16% paraformaldehyde (Electron Microscopy Sciences #15710-S) directly to the 50μL cell culture medium in each well. Plates were incubated for 30 minutes at room temperature. The well contents was aspirated using the Biotek EL406 and 50μL/well of Permeabilization / Block buffer added (Phosphate buffered saline (PBS, pH 7.5), Triton X100 0.5% (v/v), BSA 0.5% (w/v), proclin 15ppm). Plates were incubated for 30 minutes then washed 3 times with 100μL/well of PBS. Twenty-five μL/well of anti-MDM2 (rabbit polyclonal, AbCam #ab58530 diluted 1:500 in PBS, BSA 0.5% (w/v), Triton X100 0.1% (v/v)) was dispensed into each well. Plates were incubated 2 hours at room temperature then washed 4 times with 100uL/well of PBS using a Biotek EL406. Twenty-five μL/well of Alexafluor 555 conjugated anti-rabbit IgG (Life Technologies #A31572, diluted 1:1000 and Hoechst 33342 1μg/mL diluted in PBS, BSA 0.5% (w/v), Triton X100 0.1% (v/v)) was dispensed into each well. Plates were incubated for 2 hours at room temperature then washed 4 times with 100μL/well of PBS using a Biotek EL406. Fluorescence images of the samples (Channel 1: 386-23_BGRFRN_ BGRFRN (DNA); Channel 2: 549-15_BGRFRN_ BGRFRN (MDM2)) were acquired using a Cellomics XTI Arrayscan with the Bioapplication "Compartmental Analysis". Channel 1 was used to define the nuclear region. Measurements were made of "Mean_CircAvgIntCh2", which is the Alexafluor 555 fluorescence intensity (MDM2) within the nuclear region measured on a per cell basis and averaged over all the measured cells. Data analysis and EC_{50} calculation was carried out in GraphPad Prism using non-linear four parameter curve fitting.

Cathepsin-B protease assay. Cathepsin-B proteolytic activity is quantitated by an LC/MS, MRM-based detection method. Briefly, varying concentrations of experimental compound are incubated with 0.5 nM human liver Cathepsin-B (EMD Millipore, #219364) and Benzyloxycarbonyl-Arg-Arg-7-amino-4-methylcoumarin (Cbz-RR-AMC) fluorogenic substrate in buffer containing 10 mM MES pH 6.0 and 1 mM DTT. The reaction is incubated for 5 minutes at room temperature, followed by quenching with the addition of an equal volume of 2% formic acid in water. Free AMC liberated by Cathepsin-B is quantitated on a Sciex 5500 QTRAP mass spectrometer (Sciex, Framingham, MA) equipped with a Biocius Rapidfire high-throughput LC system (Agilent, Santa Clara, CA). AMC product is captured on an Agilent C18 Rapidfire cartridge, desalted with a 0.1% formic acid wash, and then eluted with 80% acetonitrile, 0.1% formic acid. Measured MRM AUC for AMC is plotted against compound concentration using GraphPad Prism (GraphPad Software, La Jolla, CA) and fitted for IC50 using four parameter fitting. MRM parameters for AMC: ESI positive mode, Q1=233.1, Q3=175.2, collision energy=16.

Caspase-3 protease assay. Caspase-3 proteolytic activity is quantitated by an LC/MS, MRM-based detection method. Briefly, varying concentrations of experimental compound are incubated with 0.025 nM of recombinant Caspase-3 (cloned and purified in-house) and 1 µM $(Z\text{-}DEVD)_2$-Rho110 fluorogenic substrate in buffer containing 25 mM Hepes, pH 7.2 and 5 mM DTT. The reaction is incubated for 20 minutes at room temperature, followed by quenching with the addition of an equal volume of 2% formic acid in water. Free Z-DEVD liberated by Caspase-3 is quantitated on a Sciex 5500 QTRAP mass spectrometer (Sciex, Framingham, MA) equipped with a Biocius Rapidfire high-throughput LC system (Agilent, Santa Clara, CA). Z-DEVD product is captured on an Agilent C18 Rapidfire cartridge, desalted with a 0.1% formic acid wash, and then eluted with 80% acetonitrile, 0.1% formic acid. Measured MRM AUC for Z-DEVD is plotted against compound concentration using GraphPad Prism (GraphPad Software, La Jolla, CA) and fitted for IC50 using four parameter fitting. MRM parameters for Z-DEVD: ESI negative mode, Q1=609.2, Q3=440.2, collision energy=-35.

USP7 aggregation analysis. Aggregation of full-length USP7 was confirmed by dynamic light scattering (DLS) using a Wyatt DynaPro Plate Reader. DLS data were acquired at 37°C, with Dynamics V7 software, with a 10 second acquisition time, 10 acquisitions per measurement in auto-attenuation mode. Compounds were present at 100µM in a buffer containing 50mM HEPES pH 7.2, 150mM NaCl, 0.01% Triton X-100, 1mM TCEP, 0.1% DMSO. Full-length USP7 was present at 1 mg/ml with all compounds except Rottlerin. Aggregate is defined here as having a hydrodynamic radius greater than 10 nm.

USP7 covalent modification and LC-MS analysis. To evaluate potential covalent modification of the proteins, full-length USP7 full-length was incubated with excess compound at room temperature overnight and covalent modification was evaluated by LC-MS using standard methods.

Cell culture and cell treatments

HCT-116 parental, USP7 null (HD R02-028), and HCT-116 p53 null (Horizon; HD 104-001) cell lines were purchased from Horizon. Normal cells were purchased from the following vendors: normal mammary cells (Life Technologies; HMEC A10565), normal osteoblasts (Lonza; CC-2538) and were cultured in the vendor-specified media. All other cell lines were obtained from Genentech's repository and were cultured in standard conditions in RPMI media containing 10% FBS (solid tumor cell lines) or 20% FBS (hematopoietic cell lines), 1% penicillin/streptomycin, and 1% L-Glutamine. All cell lines

were mycoplasma-tested and cell line identity was by STR and SNP profiling as described[26]. For studies evaluating cellular effects of compound treatments, 2,500-5,000 cells were seeded in 1-well of a 96-well plate (Corning; 3904). The following day, the media was changed from normal (10%) to low serum (0.5%) containing vehicle (DMSO) or compounds. Sixteen to 24 hours later FBS was added to each well to bring serum levels back to 10%. Cells were then allowed to grow for two additional days. All treatments were done in triplicate. For degradation rescue studies, the proteasome inhibitor bortezomib (Selleckchem) or the UAE inhibitor MLN7243 (Active Biochem) were added at 5μM 30-45 minutes prior to harvest.

Cycloheximide chase studies

MCF-7 cells were treated for a total of 7 hours with DMSO or 15μM of the indicated compounds. During the 7 hours of compound treatment, 50μM cycloheximide was added for the indicated times prior to harvest. Cell lysates were subsequently processed for western blot analysis.

Antibodies and reagents

Antibodies to the indicated proteins were purchased from the following vendors: USP7 (ab84098), USP5 (ab84695) [AbCam] or (4833) [Cell Signaling Technology]; MDM2 (Santa Cruz sc-965 or EMD Millipore 07-575); tubulin (LICOR 926-42211); tubulin-HRP (5346), actin-HRP (12620), GAPdH-HRP (8884), caspase-3 (9662), PARP (9441), total S6 (2217), phospho-S235/236-S6 (2945), Mcl-1 (4572), PIM2 (2730), cleaved-caspase-3 (2664), phospho-S112-Bad (5284), p21 (2947) [Cell Signaling Technology]; K48 polyubiquitin (Genentech); USP47 (Bethyl A301190A); UCHL1 (Invitrogen 38-1000); HA-HRP (clone HA-7) and FLAG (A8592) [Sigma]; p53 (Thermo Scientific MS738-P1); p21 (Millipore 05-655), Bad (AF819) [R&D]. Secondary antibodies were purchased from LICOR Biosciences or Jackson Immunoresearch. Full-length human USP7 was cloned into a pRK5 mammalian expression construct with a C-terminal FLAG-tag and mutations were introduced by QuickChange site-directed mutagenesis as instructed by the manufacturer (Agilent).

Western blotting, quantitation, immunoprecipitations, and deubiquitinase assays

Cells were treated with compounds as detailed above prior to lysis for Western Blot analysis. Primary antibodies were diluted 1:1000 and incubated 1 hour at room temperature or 4ºC overnight. Secondary antibodies were diluted 1:10,000 and incubated 30-60 minutes at room temperature. Blots were either imaged using a chemiluminescent reagent (Pierce) or were scanned and bands were quantified using LICOR Odyssey instrumentation and software, respectively. Immunoprecipitations with ubiquitin-specific antibodies were performed as described[27] and immunoprecipitates and corresponding cell lysates were analyzed by immunoblot analysis as previously described[27] using the antibodies detailed above. For *in vitro* deubiquitination of PIM2, ubiquitinated PIM2 was captured by treating MV-4-11 cells for 45 minutes with the proteasome inhibitor MG-132 (SelleckChem). Cells were washed, lysed in a 6M urea lysis buffer, and anti-K48 polyubiquitin immunoprecipirates were washed and deubiquitinated with 250nM wild-type or C223S full length recombinant USP7 following a similar protocol as described[27].

Cell viability studies

IncuCyte live cell analysis. One day after cell seeding, 10% cell culture media was changed to 0.5% serum media containing 2μM CellEvent Caspase 3/7 reagent (Life Technologies; C10423) and compounds. The plates were placed in an IncuCyte live cell

imager and scanning was started within 15-20 minutes. Images were taken every 2 hours for 68-72 hours, using a 10X objective. Phase contrast was used to measure cell confluency/density while green fluorescence was used to measure caspase activity. The images were analyzed using IncuCyte software (Basic Analysis parameters) and a ratio of caspase activity to cell density/count was determined. For combination experiments, MCF7 cells were treated with 15μM USP7 inhibitors or 0.1μM doxorubicin alone or in combination. Similarly, U2OS cells were treated with 15μM USP7 inhibitors or 1μM cisplatin alone or in combination.

CellTiter-Glo assays (Extended Data Figures 1d, 1e, 2b) . Seventy-two hours after compounds were added, CellTiter-Glo (CTG, Promega) reagent was added following the vendor protocol. Three times more USP7 null HCT-116 cells were plated per well than wild-type HCT-116 cells (7,500 vs. 2,500 cells), given the slower proliferation of USP7 null cells. For studies with multiple myeloma cell lines, cells were seeded in 0.5% serum and treated with compounds immediately. Twenty-four hours later, serum was added back to normal levels. CTG assay was done 24 hours later, i.e. 48 hours after compounds were added instead of 72 hours later for the adherent cells.

Tumor cell line panel viability studies. GNE-6640 and GNE-6641 were profiled across 441 cell lines as previously described[28]. In brief, compounds were screened in 9-point dose response using a 3-fold dilution. Cells were seeded into 384 well plates 24 hours prior to compound addition. Cells were then incubated with compound for 72 hours or 120 hours before assaying viability (CellTiter-Glo, Promega). Assays were performed in biological triplicate. Cells were incubated (37 °C, 5% CO_2) in RPMI-1640, 2.5% FBS (72 hour assay) or 5% FBS (120 hour assay), and 2 mM glutamine throughout the assay. The reported IC_{50} and mean viability metrics are as follows: IC_{50} is the dose at which the estimated inhibition is 50% relative to untreated wells (i.e. absolute IC_{50}). The mean viability statistic is the arithmetic average of the fitted viabilities at each tested dose. Mean viability is equivalent to the area under the log-dose/viability curve divided by the total number of tested doses, and is thus on an interpretable percentage scale.

Exome-Seq pipeline. FASTQ reads were aligned to the human reference genome (GRCh38) using GSNAP (PMID:20147302, 27008021) version '2013-10-10' using the following parameters: -M 2 -n 10 -B 2 -i 1 --pairmax-dna=1000 --terminal-threshold=1000 --gmap-mode=none --clip-overlap. Duplicate reads in the resulting BAM file were marked using PicardTools, and indels realigned using the GATK IndelRealigner tool. Variations were called using the Bioconductor package VariantTools version 1.9.4. using default options except for two exceptions: 1) no variants were called in repeat regions as defined by the annotation Dust, Satellite repeats, and Tandem repeats in EnsEMBL 77 and 2) the avgNborCount post filter was configured using all SNPs from dbSNP version 138.

Mean viability analysis. We determined genomic and non-genomic features that were associated with differences in compound sensitivity. Mean viability is calculated as the arithmetic average of the fitted viabilities at each tested dose, as previously described[28]. For this analysis, normalized mean viabilities of GNE-6640 and GNE-6446 were determined by normalizing the mean viabilities of each compound by the mean viability of GNE-6641. Cancer type, cell histology, loss-of-function and hotspot mutations were assessed for statistical association with changes in normalized mean viability. Loss-of-function protein coding mutations include: 1) insertions, 2) deletions, and 3) substitutions resulting in predicted truncating, splice site, translational start site, or non-stop mutations

present in greater than 90% of reads sequenced (> 90% variant allele frequency). Hotspot missense mutations assessed here were previously reported as significantly recurrent mutations in a pan-cancer analysis (PMID: 26619011). Features present in at least 3 cell lines were assessed. Statistical significance was determined using two-sided Student's t-Test. Q-values were determined by correcting resulting p-values for multiple hypothesis testing using Benjamini and Hochberg.

Primary combination screen. A compound library comprising 589 compounds arrayed in 9 point dose response was screened in the absence or presence of fixed doses of GNE-6776 (0 nm, 125 nM, 250 nM, 500 nM, 1000nM, and 2000 nM) or GNE-6640 (400 nM). Briefly, 5,000 EOL-1 cells were seeded into 384 well plates, and compound was added 24 h later. Cell viability was determined 120 h post-compound addition (CellTiter Glo). Curves were fitted, and both IC_{50} and mean viability metrics were calculated. The IC_{50} is the dose at which inhibition is 50% relative to untreated wells. The mean viability is the average of the fitted viabilities at each tested dose. Mean viability is equivalent to the area under the log-dose/viability curve divided by the total number of tested doses. Mean viability values were used for the analysis described in Extended Data Figure 6d. All data were fitted using Genedata Screener (GDS) software.

Primary combination screen analysis. We determined normalized mean viabilities in the EOL-1 cell line for 574 compounds with a known protein or mechanistic target either in the presence of DMSO or increasing concentrations of GNE-6776 (100 nM, 250 nM, 500 nM, 1000 nM or 2000 nM) or 400 nM of GNE-6640. For each compound we assessed the difference in mean viability between USP7-inhibitor treatment versus the DMSO treatment. For targets targeted by 3 or more compounds we calculated the enrichment of high mean viability difference for each concentration of USP7 inhibitor by using the Wilcoxon rank sum test. For visualization purposes we combine the results of all concentrations by taking the mean of the -log10 transformed p-values for each target.

Bliss Analysis. PIM inhibitors were tested in a 9-point dose response matrix with GNE-6676 in the same manner as described for the compound library screen. Inhibitors were screened in 3-fold dilution using a top concentration of 10 μM for PIM inhibitors and 20 μM for GNE-6676. Bliss calculations were performed in GDS.

Deubiquitinase selectivity analysis

Recombinant deubiquitinase di-ubiquitin mass spec cleavage assay. The MALDI-TOF DUB assay was performed using the indicated concentrations of recombinant deubiquitinases, di-ubiquitin substrates, and USP7 inhibitor compounds as previously described[4]. The inhibition efficiency for GNE-6640 and GNE-6776 against the UCHL family members was monitored on an alternative substrate, Ub-Ube2W (Ub-E2), since UCHL1, -3, and -5 and BAP1 are inactive against ubiquitin dimers of all linkage types[29].

Endogenous deubiquitinase activity-based probe assay. 293T cells at 80% confluency were harvested by rinsing the plate once with 10 mL PBS followed by scraping. Cells were cleared by spinning them for 3 min at 350 g at 4°C and cell pellets were flash-frozen in liquid nitrogen and stored at –80 °C until lysis. Frozen pellets were lysed by quickly re-thawing them in Buffer A (50 mM Tris-HCl pH 7.5, 250 mM Sucrose, 1x Phosphatase STOP, 2.5 mM TCEP, 2 mM ATP, 50 μM phenylmethylsulfonyl fluoride (PMSF), 120 mM NaCl, 5 mM $MgCl_2$) and the lysate was cleared by centrifugation by spinning at 18,000 g at 4°C. The protein concentration was adjusted to 5 mg/mL, and 5 mg (1 mL) each of this cell lysate was incubated with indicated compounds or DMSO at

the indicated concentrations for 20 min at 900 rpm, 25 °C in a Thermomixer® (Eppendorf AG). After compound incubation, 300 ng recombinant viral DUB was added and lysates were incubated with 6.6 µg/mL of the indicated activity-based DUB probe, 1250 rpm, at 25°C, for the indicated times. The reaction was terminated by adding a 20% SDS solution to a final concentration of 0.4% for at least 30 min at room temperature, rotating. Subsequently, the lysate was diluted 10x with Buffer B (50mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM EDTA, protease inhibitor cocktail EDTA-free (Roche, Mannheim, Germany), 50 µM PMSF, 0.5 % NP-40) into 15 mL conical tubes. ~ 120 µL slurry of pre-equilibrated anti-HA affinity matrix (Roche, Mannheim, Germany) was added and HA-tagged proteins were immunopurified with this matrix by rotating the samples overnight at 4 °C. The beads were then washed 1 mL of ice-cold Buffer B (3x), Buffer B without NP-40 (1x), and TEAB (15 mM Triethylammoniabicarbonate, pH 8.5) (3x). Spins between washes were performed at 2000 g at 4°C. To elute the immunopurified material 330 µL of Buffer C (1 mg/mL HA peptide (Thermo Scientific, Waltham, MA), 15 mM TEAB, 0.02 % Rapigest® (Waters, Milford, MA)) was added and the samples were incubated at 37°C for 30 min at 1100 rpm shaking in a Thermomixer® (Eppendorf AG). The eluted material was cleared from the beads by spinning at 2600 g. The cleared material was stored at –80 °C until mass spectrometry preparation and analysis. For DUB identifications the eluted proteins were digested with trypsin using filter-aided sample preparation (FASP). Eluents were added to Microcon-30K filtration devices (Millipore, Billerica, MA) and briefly washed by 0.2 mL of 8 M urea in 200 mM TEAB, pH 8.5. In between each step, liquid was cleared by centrifugation at 14,000 xg, except if noted. Proteins were reduced by Dithiothreitol (DTT) for 20 minutes at 60°C and subsequently alkylated by iodoacetamide (IAA) for 15 minutes in the dark. The membrane was further washed by 8 M urea once and 200 mM TEAB three times. Trypsin (Promega, Madison, WI) was added at 1:40 enzyme/substrate ratio. Devices were briefly centrifuged at 100 xg for 30 seconds and incubated overnight at 37°C. Tryptic peptides were recovered by centrifugation at 14,000 xg for 4 minutes. An additional 60 µL of 200 mM TEAB, pH 8.5 was added to the devices and centrifuged as an additional elution step. This eluent was combined with the previous for further processing. Ten percent of the eluents were dried in speedvac and desalted with a C18 STAGE tip (Proxeon, Thermo Fisher) before injected onto the mass spectrometer for LC-MS/MS acquisition.

Deubiquitinase identification by mass spectrometry. Peptides were loaded onto an Acuity UPLC® BEH130 C18 column (1.7 µm, 12Å, 100 µm x 100 mm) at a flow rate of 1.5 µL/min in solvent A (98% water/2% MeCN/0.1% formic acid) using a NanoAcquity UPLC system (Waters, Milford, MA). Separation was achieved with a linear gradient of 2% Solvent B (98% MeCN/2% water/0.1% formic acid) to 25% Solvent B over 45 minutes. Eluted peptides were injected onto an Orbitrap Elite mass spectrometer (Thermo Fisher, San Jose, CA) using an Advance CaptiveSpray source (Bruker, Auburn, CA) at a voltage of 1.2kV. Full MS scans were collected in the orbitrap at 60,000 resolution and the top 15 most abundant ions were selected in a data dependent mode and fragmented with CID and ms/ms were collected in the ion trap. MS/MS spectra were searched using Mascot (v.2.3.02) against a human proteome database (Uniprot Dec. 2011) with known contaminants along with all decoy sequences. Search parameters included trypsin cleavage allowing up to 2 missed cleavage events, a precursor ion tolerance of 50 ppm, and a fragment ion tolerance of 0.8 Da. Searches also permitted variable modifications of methionine oxidation (+15.9949 Da), and two cysteine modifications of either (+57.0215) for carbamidomethylation or (+192.0569 Da) for reacted DUB probe remnant, Peptide spectra matches were filtered with a false

discovery rate (FDR) of 5% on the peptide level and subsequently at 2% on the protein level. Each identified peptide-spectrum match (PSM) was quantified using the area under curve and same quantification event was extended to other runs where such peptide was not identified by an ms/ms spectrum, based on exact precursor m/z and retention time matching.

Statistical analysis of deubiquitinase activity data. Detected and quantified area-under-curve (AUC) measures from label-free mass spectrometry were logarithmically (base 2) rescaled. Changes between inhibitor- and control-treated lysate were compared via a linear mixed effects model with fixed effects for treatments and random effects for peptide species and mass-spec run. Computations were performed in R, version 3.3.0 using the R package *lme4*, with error degrees of freedom estimated via the Kenward-Roger method using the R package *pbkrtest*. For each protein, $\log_2$ fold change in activity was normalized to that observed in a spiked-in viral DUB or BSA control protein. Statistical significance of the observed "inhibitor-vs-control" differences in $\log_2$ fold change between each DUB and the control protein was then assessed via a t-statistic with Satterthwaite-estimated degrees of freedom, and corrected for multiplicity via the Benjamini-Hochberg False-Discovery Rate method (FDR) set at 0.10.

Animal use and care

All animal work followed the recommendations of the Guide for Care and Use of Laboratory Animals with respect to restraint, husbandry, surgical procedures, feed and fluid regulation, and veterinary care. The animal care and use program at Genentech is accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International (AAALAC), which assures compliance with accepted standards for the care and use of laboratory animals. Studies were tailored to minimize the number of animals used, yet sufficient numbers to address any variability in drug exposure or biomarker response. Due to the need to monitor potential adverse effects with first in-life assessment of novel compounds, no study was conducted under blinded conditions.

DMPK analysis

In vitro DMPK studies were performed using standard protocols. GNE-6776 was formulated as a suspension in 0.5% methylcellulose/0.2% Tween-80 (MCT) and was administered at 200 mg/kg by oral gavage to female *C.B-17 SCID* mice, age 12 – 16 weeks (Charles River Labs; n=3 per time point). At 0.5, 1, 2, 4, 8 and 24 hours post-dose, blood samples were collected by terminal cardiac puncture into anticoagulant tubes (EDTA). Clarified plasma was then transferred to a fresh tube and snap frozen. GNE-6776 plasma concentrations were determined by LC/MS-MS.

In vivo *pharmacodynamic response*

For EOL1 acute myeloid leukemia xenograft studies, immunodeficient *C.B-17 SCID* mice (Charles River Labs), aged 12 - 16 weeks, were inoculated subcutaneously on the right flank with five million cells in a 50:50 suspension of HBSS:Matrigel (BD Biosciences; 100 µL). When tumor volumes reached between ~285-500 mm^3, mice were distributed into volume-matched cohorts (n=4). For MCF7 breast cancer xenograft studies, immunodeficient *nu/nu* mice (Charles River Labs), aged 6-8 weeks, were implanted with 0.36 mg estrogen pellets (Innovative Research of America) via trochar 1-3 days prior to tumor cell inoculation. Ten million MCF7-Ser cells, an *in vivo*-optimized MCF7 variant, were injected orthotopically into the 2/3 mammary fat pad of each mouse in a 50:50 suspension of HBSS:Matrigel (BD Biosciences) in a total volume of 100 µL. When tumor volumes reached between ~285-450 mm^3, mice were distributed into volume-matched

cohorts (n=4). GNE-6776 was formulated as a suspension in 0.5% methylcellulose/0.2% Tween-80 (MCT) and was administered at 200 mg/kg by oral gavage at zero and four hours. MCT control or GNE-6776-treated samples were collected at eight hours following the first dose and excised tumors were flash-frozen on dry ice. Tumors were lysed in RIPA buffer containing protease inhibitors (Roche) and 300 mM NaCl using a Qiagen Tissuelyser. Samples were incubated on ice for 15 minutes and then centrifuged at 20,000 x g at 4˚C for 10 minutes. Protein levels in clarified lysates were quantified using a Pierce BCA assay kit and concentrations were normalized with sample buffer. Samples were run on gels and proteins were transferred to membranes and western blotted as described above.

In vivo efficacy study

For EOL1 acute myeloid leukemia xenograft studies, immunodeficient *C.B-17 SCID* mice (Charles River Labs), aged 12 - 16 weeks, were inoculated subcutaneously on the right flank with five million cells in a 50:50 suspension of HBSS:Matrigel (BD Biosciences; 100 µL). When tumors became established (150 – 300 mm^3), mice were distributed into tumor volume-matched cohorts (n=7, mean tumor volume ~250 mm^3). GNE-6776 was formulated as a suspension in 0.5% methylcellulose/0.2% Tween-80 (MCT) and was administered at 100 or 200 mg/kg by oral gavage on a once or twice daily schedule. Tumor volume measurements, body weight and condition data was collected 2-3 times per week.

USP7 enzymatic analysis

Michaelis-Menten kinetic measurements with full-length USP7 were carried out using 1 nM USP7 with a series of ubiquitin-AMC substrate titrations with at least three technical replicates. Initial rate of substrate hydrolysis was determined using the Magellan software on a Tecan Safire2 plate reader and kinetic parameters modeled using nonlinear regression analysis with GraphPad Prism software. Standard error was calculated from multiple experimental replicates. For studies using the USP7 D305/E308 mutant, samples were reacted in a buffer consisting of 50 mM HEPES (pH 7.5), 100 mM NaCl, 2.5 mM Dithiothreitol, 0.1% (w/v) bovine gamma globulin (Sigma cat # G5009-25G). The starting substrate concentration of Ubiquitin-Rho110 (Boston Biochem cat # U-555) used for the Michaelis-Menten analysis was 100 µM serial diluted to 781 nM. Reactions were carried out for 1 hour at room temperature with a final enzyme concentration of 100 nM (three independent experiments, see symbols in plot), in a black 100 µL volume 96 well half area plates (Corning cat # 3993). The enzymatic activity was calculated by fitting the data using the initial velocity using the linear V_0 values measured by analyzing the fluorescence signal of cleaved Rho-110 using excitation at 485 nm and emission at 535 nm.

X-Ray crystallography

Crystals were grown by the hanging-drop method by mixing the USP7 catalytic domain (residues 208–554) at 15 mg/ml with an equal volume of reservoir solution containing 100 mM Tris,-HCl, pH 7.0, and 20% PEG1000 (v/v). Co-structures with compounds were obtained by soaking crystals with 1 mM of compound overnight. Crystals were cryoprotected with reservoir solution supplemented with 20% glycerol (v/v) and flash frozen in liquid nitrogen. Data collection and refinement statistics are detailed in Extended Data Fig. 7d.

Generation of isotopically-labeled di-ubiquitins

Ubiquitin was cloned into a peT3a vector and transformed into the auxotrophic strains RF2 and ML2 (a kind gift from Robert Gennis and Toshio Iwasaki) [30] and expressed with the following modifications: Bacterial cultures were grown at 37°C in M9 media supplemented with NH_4Cl (2 g/L), C_6-glucose (4 g/L). Strain RF2 was additionally supplemented with 6mM ^{15}N L-Thr (Cambridge Isoptopics Cat#NLM-142-PK) and ML2 were supplemented with 6mM ^{15}N L-Leu (Cambridge Isoptopics Cat#NLM-142-1) and 6 mM unlabeled Ile and Tyr (Sigma-Aldrich Cat# I2752 and Cat# T3754). After reaching an OD of 0.6 the cells were induced with 1 mM IPTG and further grown for 6 hours. Protein purification was performed at room temperature. Cells were harvested and lysed in lysis buffer (50 mM HEPES 7.0). The cleared lysate was subjected to affinity chromatography using DEAE sepharose Fast Flow (GE Cat# 17-0709-01). Ubiquitin was collected in the flow through and dialyzed overnight into NaOAc pH 4.5. Dialyzed material was clarified by centrifugation at 35K and ubiquitin was subjected to ion exchange chromatography (IEX) using a MonoS column (GE Healthcare, Cat# 17-5169-01). The following enzymes were obtained from Boston Biochem Cambridge, UBE1 (Cat# E-305), UBE2K and UBE2N/UBE2V1 complex (Cat# E2-602 and Cat# E2-664) respectively. K63- and K48-linked di-ubiquitin chains were generated and purified as follows: In separate reactions incubating 250 nM E1 enzyme, 5 µM UBE2K (K48 linked) or 5 µM UBE2N/UBE2V1 complex (K63 linked) with equal molar ratios of 1 mM ubiquitin and 1 mM ubiquitin, 10 mM ATP, 50 mM HEPES (pH 8.0), 10 mM $MgCl_2$ in a 10 mL reaction at 37 °C. After 2 hours, the reaction was acidified with 2 mL of 17.4 M Glacial Acetic Acid. Obtained di-ubiquitins (K63- or K48-linked) were purified by cation exchange using a MonoS column (GE Healthcare, Cat# 17-5169-01). All purified di-ubiquitin chains were buffer exchanged into 1x PBS buffer and proteins were flash frozen in liquid nitrogen prior to storage at −80°C.

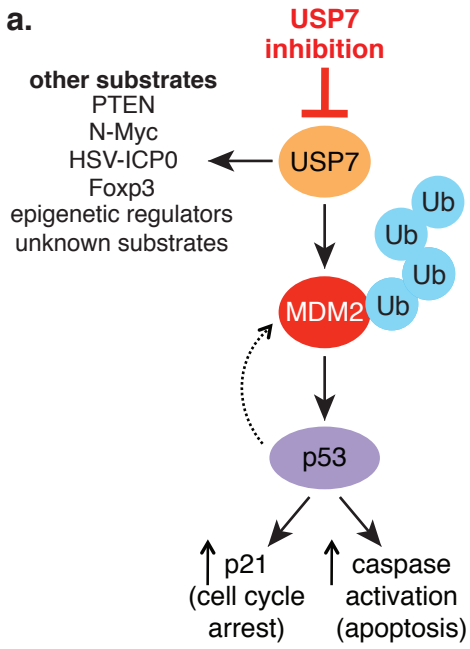
Generation of 5-TAMRA-peptide/tetra-ubiquitin conjugates

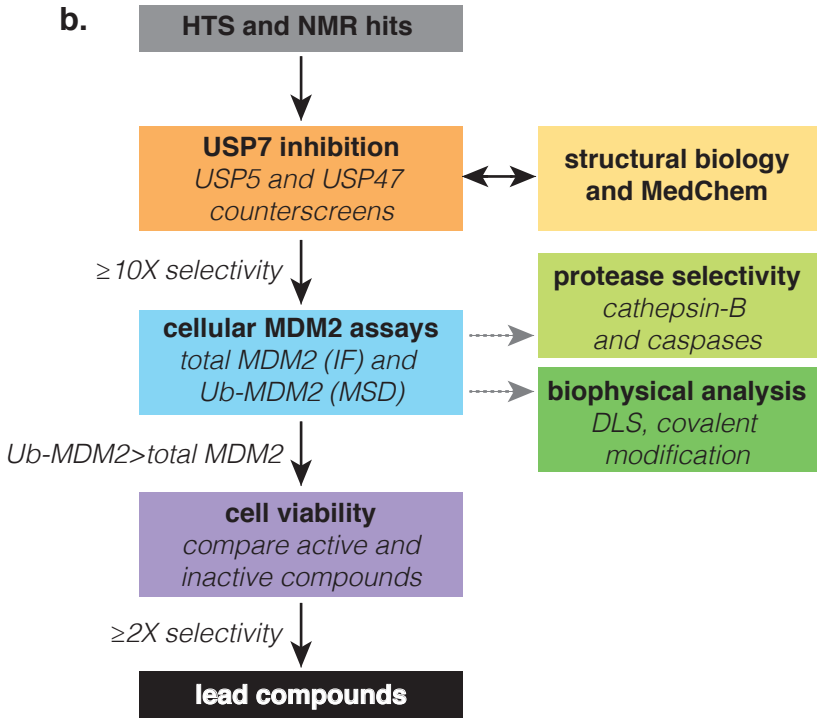
K63- and K48-linked tetra-ubiquitin chains were obtained from Boston Biochem Cambridge, K63-linked (Cat# UC-310B), K48-linked (Cat# UC-210B). The 5-TAMRA (5-Carboxytetramethylrhodamine) peptide was generated by CPC-Scientific consisting of the sequence 5-TAMRA-YPYDVPDYAIREIVSRNKRRYQEDG[20]. K63 or K48 tetra-ubiquitin chains were conjugated to the peptide as follows: (1) Generation of tetra-ubiquitin-MESNA; incubating 250 nM E1, 10 mM $MgCl_2$, 10 mM MgATP, 1 mM tetra-ubiquitin, 100 mM MESNA (Sigma Aldrich, Cat# 63705), in 20 mM Na_2HPO_4 at pH 8.0 at 37˚C overnight. Dialyzed into 0.4% TFA and tetra-ubiquitin-MESNA was lyophilized. (2) Lyophilized tetra-ubiquitin was dissolved in DMSO at a concentration of 0.5 mM and 2 mg of peptide were added until all components were dissolved, reaction volume 1 mL. The reaction was initiated by adding (final concentrations) NHS 27.5 mM, $AgNO_3$ 3.3 mM, and 22 µl DIPEA and incubated at RT overnight. The reaction was diluted 10X with ddH2O and desalted into PBS pH 7.5. Non-conjugated peptide was removed by size exclusion (SEC) chromatography. In a second step non-conjugated tetra-ubiquitin was removed by IEX chromatography as described above and buffer exchanged into 1x PBS (pH 7.5). The concentration of the purified final conjugate was determined by absorbance using an extinction coefficient for 5-TAMRA at 80,000 $cm^{-1}M^{-1}$.

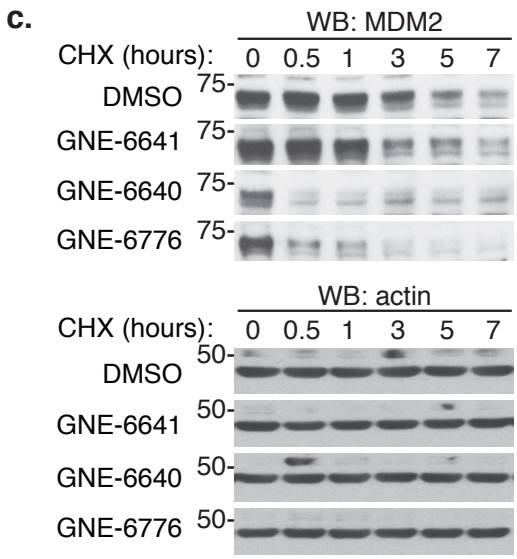
TAMRA-peptide/tetra-ubiquitin conjugate depolymerization studies

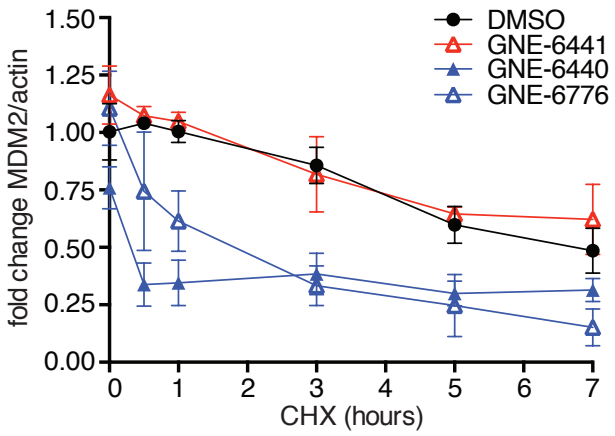
For depolymerization assays, 100 nM USP7 (cat# E-519, lot#09939314) or 100 nM USP7 catalytic domain N-terminal His-tag (Genentech, Hs_USP7.K208-K554) were diluted in 1x PBS buffer (pH 7.5) containing 5 mM DTT to generate 10x stock solutions in respect to the final concentration and preincubated at room temperature for 10 minutes. In a 90 µL reaction, 9 µg (2.7 µM) of 5-TAMRA-peptide/tetra-ubiquitin (K63- or K48-

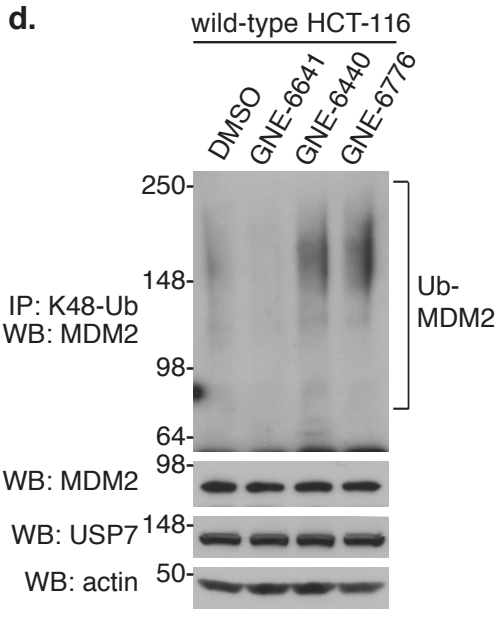
linked) was mixed with 9 µL of diluted enzyme in 1 x PBS buffer (pH 7.5). Aliquots of 10 µL of the reaction were mixed with 4 µL 2x SDS loading buffer at the time points indicated to stop the reaction. Samples (14 µL) were subjected to SDS gel-electrophoresis using precast BioRad Criterion TGX 10-20% gels (cat# 5671114). Fluorescence was analyzed using the FluorChem imager from Protein Simple according to the user manual. Densitometry values were analyzed using the software AlphaViewSA, ProteinSimple.

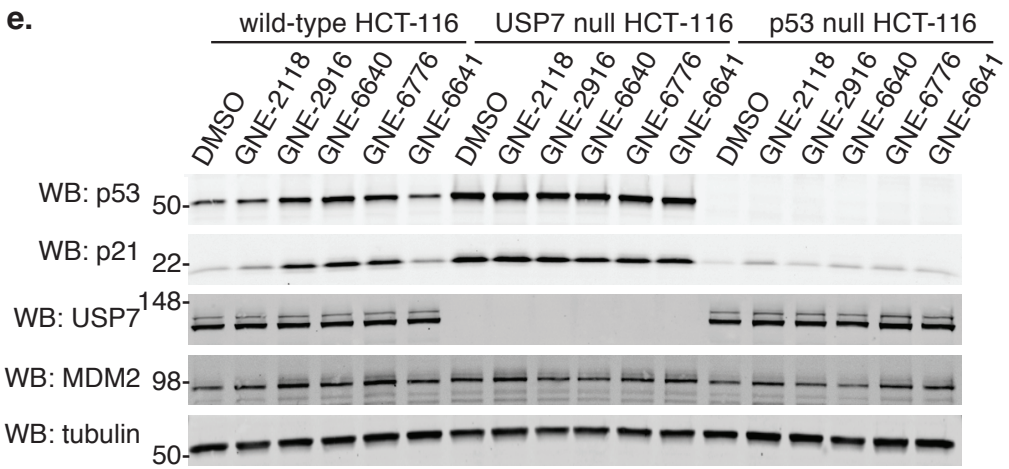

a.
USP7 inhibition
other substrates
PTEN
N-Myc
HSV-ICP0
Foxp3
epigenetic regulators
unknown substrates
USP7
Ub
Ub
Ub
Ub
MDM2
p53
p21
(cell cycle
arrest)
caspase
activation
(apoptosis)

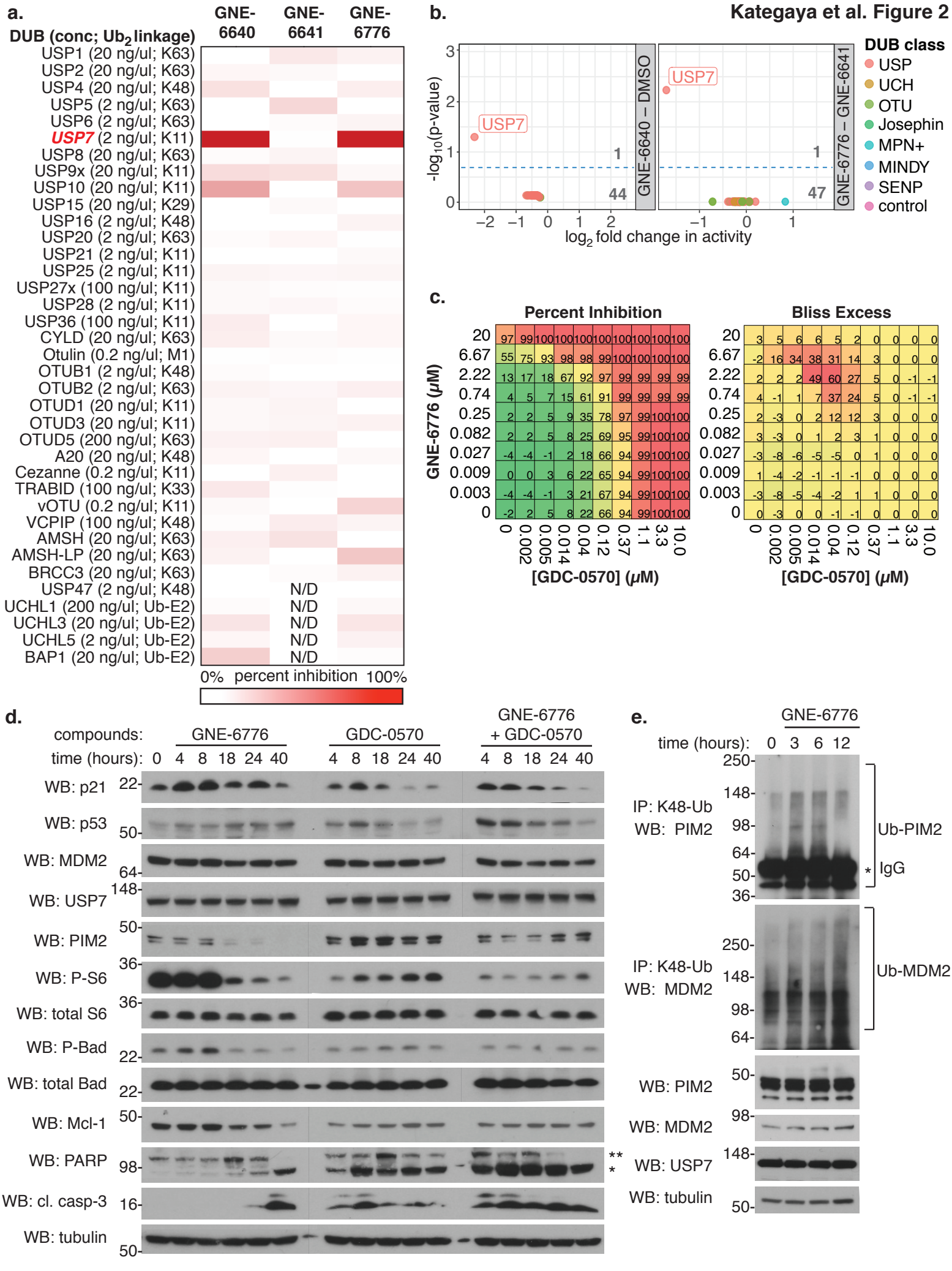

b.
HTS and NMR hits
USP7 inhibition
USP5 and USP47
counterscreens
structural biology
and MedChem
≥10X selectivity
cellular MDM2 assays
total MDM2 (IF) and
Ub-MDM2 (MSD)
protease selectivity
cathepsin-B
and caspases
biophysical analysis
DLS, covalent
modification
Ub-MDM2>total MDM2
cell viability
compare active and
inactive compounds
≥2X selectivity
lead compounds

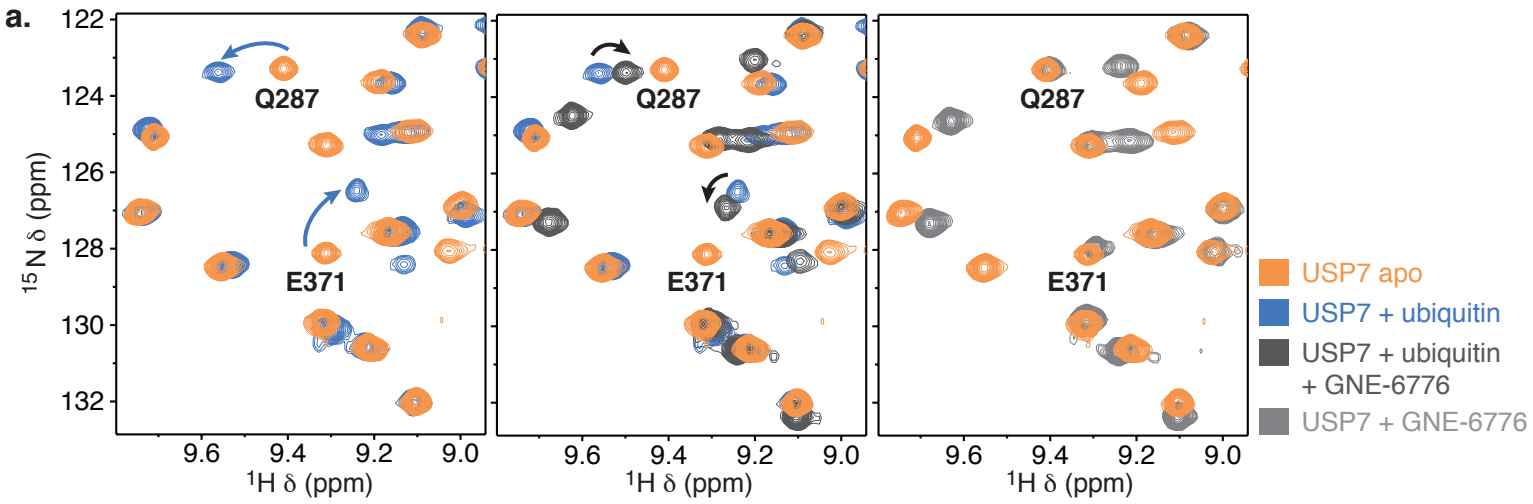

c.
WB: MDM2
CHX (hours): 0 0.5 1 3 5 7
DMSO
GNE-6641
GNE-6640
GNE-6776
75-
WB: actin
CHX (hours): 0 0.5 1 3 5 7
DMSO
GNE-6641
GNE-6640
GNE-6776
50-

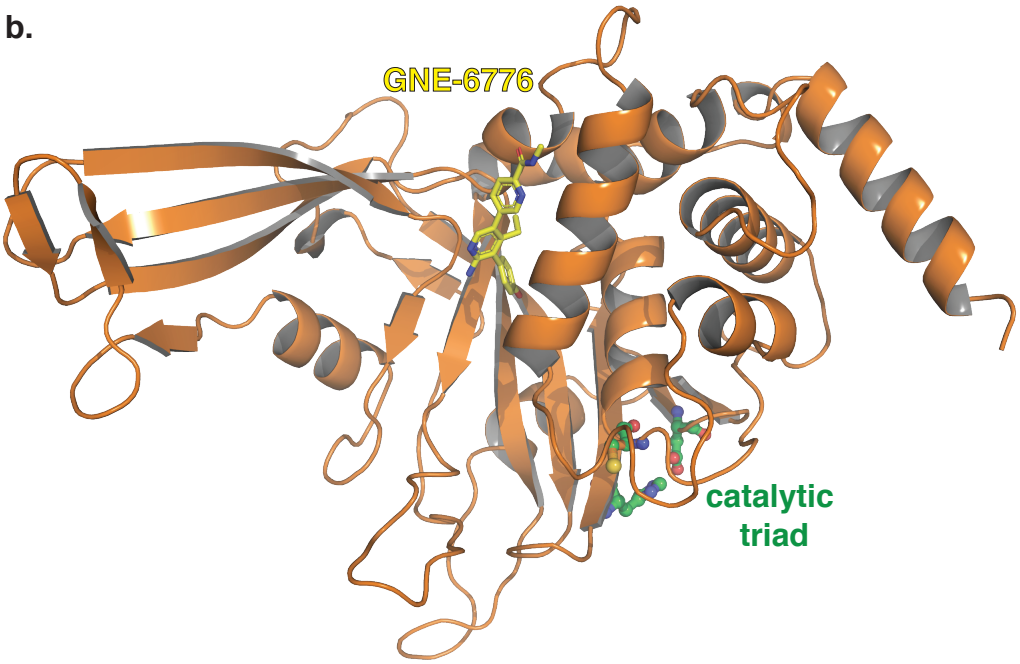

DMSO
GNE-6441
GNE-6440
GNE-6776
fold change MDM2/actin
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0 1 2 3 4 5 6 7
CHX (hours)

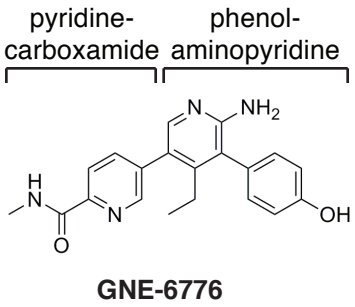

d.
wild-type HCT-116
DMSO
GNE-6641
GNE-6440
GNE-6776
250-
148-
98-
64-
IP: K48-Ub
WB: MDM2
Ub-
MDM2
WB: MDM2 98-
WB: USP7 148-
WB: actin 50-

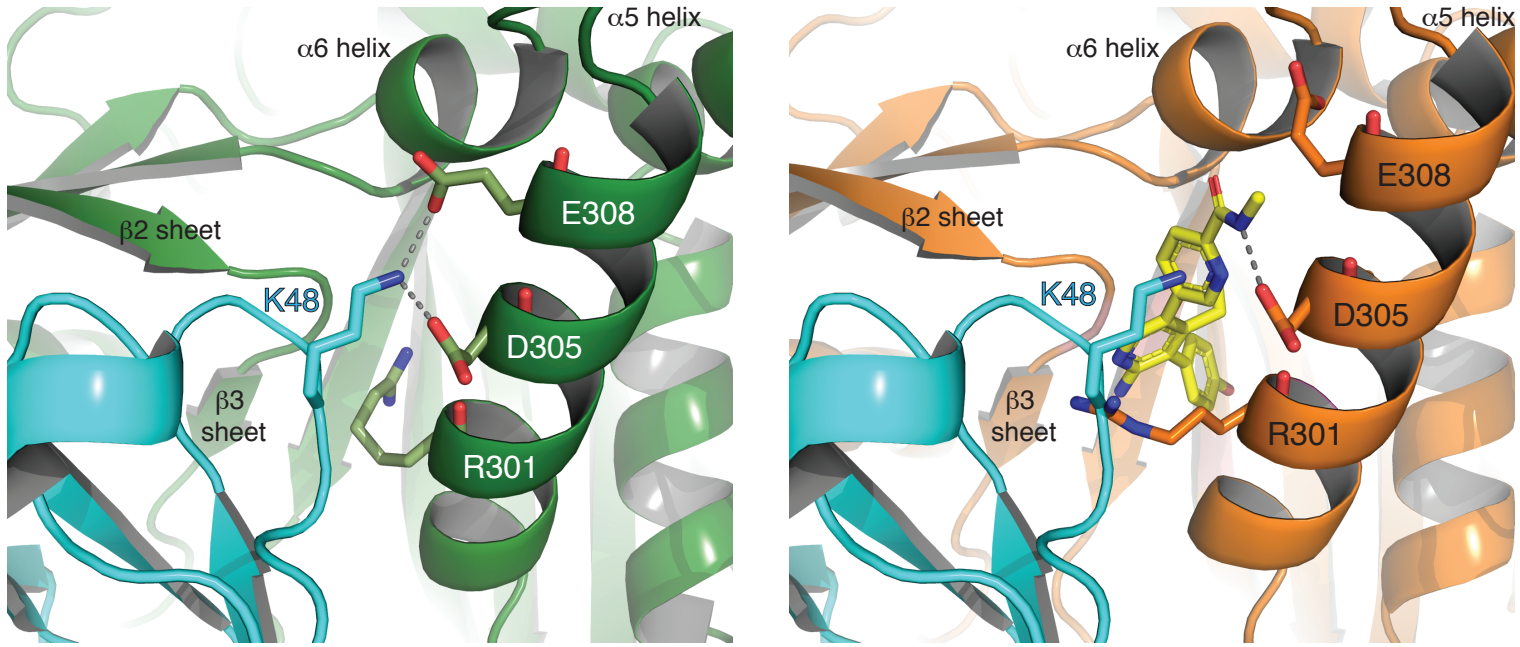

e.
wild-type HCT-116
USP7 null HCT-116
p53 null HCT-116
DMSO
GNE-2118
GNE-2916
GNE-6640
GNE-6776
GNE-6641
WB: p53 50-
WB: p21 22-
WB: USP7 148-
WB: MDM2 98-
WB: tubulin 50-

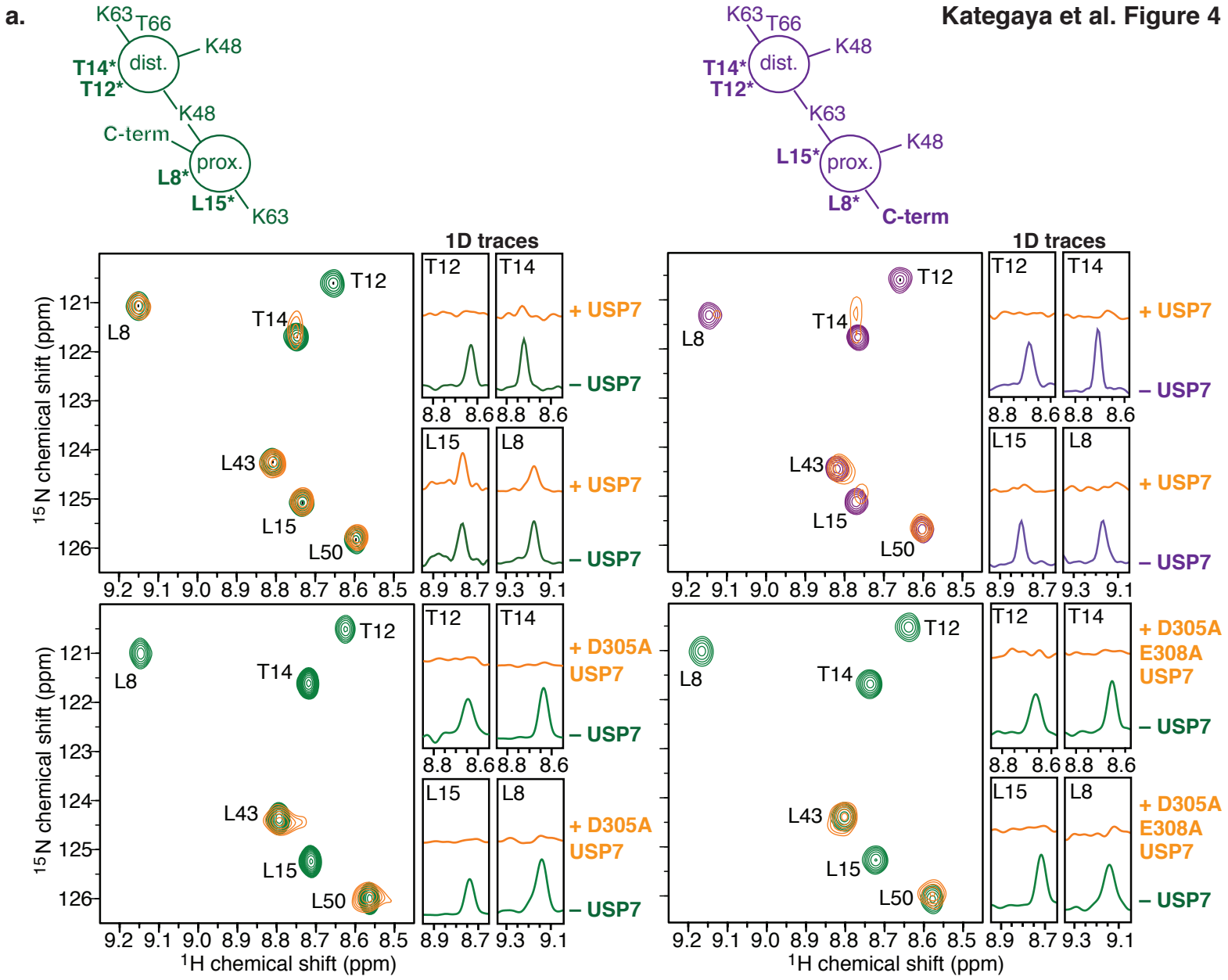

a.
GNE-6640
GNE-6641
GNE-6776
DUB (conc; Ub2 linkage)
USP1 (20 ng/ul; K63)
USP2 (20 ng/ul; K63)
USP4 (20 ng/ul; K48)
USP5 (2 ng/ul; K63)
USP6 (2 ng/ul; K63)
USP7 (2 ng/ul; K11)
USP8 (20 ng/ul; K63)
USP9x (20 ng/ul; K11)
USP10 (20 ng/ul; K11)
USP15 (20 ng/ul; K29)
USP16 (2 ng/ul; K48)
USP20 (2 ng/ul; K63)
USP21 (2 ng/ul; K11)
USP25 (2 ng/ul; K11)
USP27x (100 ng/ul; K11)
USP28 (2 ng/ul; K11)
USP36 (100 ng/ul; K11)
CYLD (20 ng/ul; K63)
Otulin (0.2 ng/ul; M1)
OTUB1 (2 ng/ul; K48)
OTUB2 (2 ng/ul; K63)
OTUD1 (20 ng/ul; K11)
OTUD3 (20 ng/ul; K11)
OTUD5 (200 ng/ul; K63)
A20 (20 ng/ul; K48)
Cezanne (0.2 ng/ul; K11)
TRABID (100 ng/ul; K33)
vOTU (0.2 ng/ul; K11)
VCPIP (100 ng/ul; K48)
AMSH (20 ng/ul; K63)
AMSH-LP (20 ng/ul; K63)
BRCC3 (20 ng/ul; K63)
USP47 (2 ng/ul; K48)
UCHL1 (200 ng/ul; Ub-E2)
UCHL3 (20 ng/ul; Ub-E2)
UCHL5 (2 ng/ul; Ub-E2)
BAP1 (20 ng/ul; Ub-E2)
N/D
0% percent inhibition 100%
b.
Kategaya et al. Figure 2
-log10(p-value)
USP7
USP7
1
1
44
47
GNE-6640 – DMSO
GNE-6776 – GNE-6641
log2 fold change in activity
DUB class
USP
UCH
OTU
Josephin
MPN+
MINDY
SENP
control
c.
Percent Inhibition
Bliss Excess
GNE-6776 (μM)
20
6.67
2.22
0.74
0.25
0.082
0.027
0.009
0.003
0
0
0.002
0.005
0.014
0.04
0.12
0.37
1.1
3.3
10.0
[GDC-0570] (μM)
d.
compounds:
GNE-6776
GDC-0570
GNE-6776 + GDC-0570
time (hours): 0 4 8 18 24 40
WB: p21
WB: p53
WB: MDM2
WB: USP7
WB: PIM2
WB: P-S6
WB: total S6
WB: P-Bad
WB: total Bad
WB: Mcl-1
WB: PARP
WB: cl. casp-3
WB: tubulin
22-
50-
64-
148-
50-
36-
36-
22-
22-
50-
98-
16-
50-
**
*
e.
GNE-6776
time (hours): 0 3 6 12
IP: K48-Ub
WB: PIM2
IP: K48-Ub
WB: MDM2
WB: PIM2
WB: MDM2
WB: USP7
WB: tubulin
250-
148-
98-
64-
50-
36-
Ub-PIM2
IgG
Ub-MDM2
*

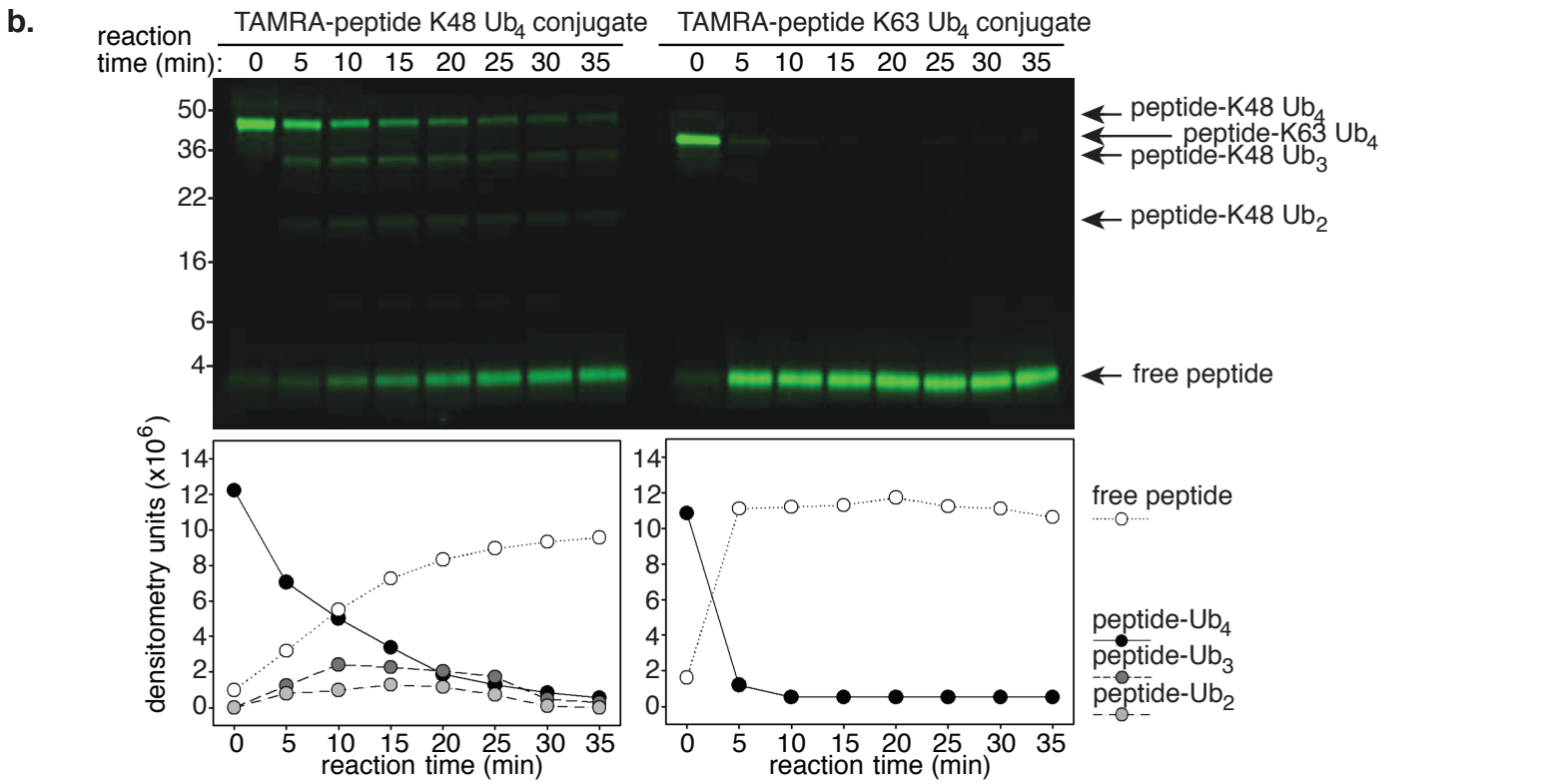

a.
Q287
E371
^{15}N δ (ppm)
^{1}H δ (ppm)
USP7 apo
USP7 + ubiquitin
USP7 + ubiquitin + GNE-6776
USP7 + GNE-6776


b.
GNE-6776
catalytic triad


pyridine-carboxamide
phenol-aminopyridine
GNE-6776

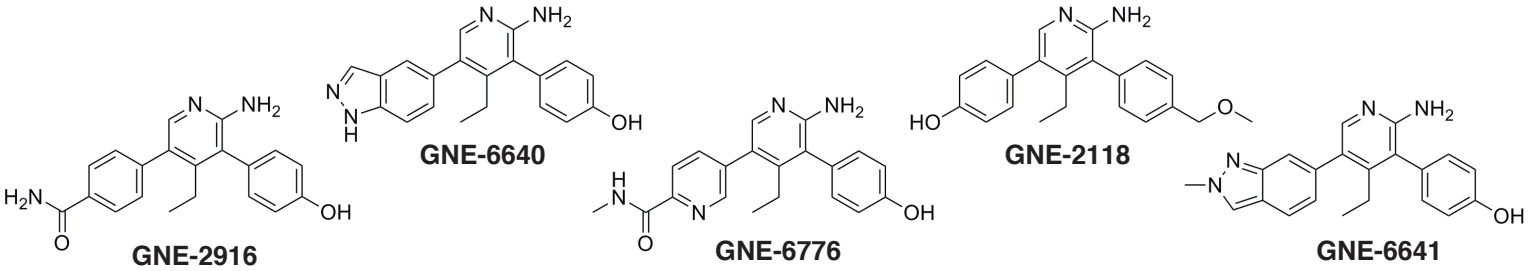

c.
α6 helix
α5 helix
β2 sheet
K48
E308
D305
β3 sheet
R301


a.
K63
T66
dist.
K48
T14*
T12*
K48
C-term
prox.
L8*
L15*
K63
K63
T66
dist.
K48
T14*
T12*
K63
L15*
prox.
K48
L8*
C-term
1D traces
1D traces
T12
T14
L8
L43
L15
L50
+ USP7
– USP7
+ D305A USP7
+ D305A E308A USP7
15N chemical shift (ppm)
1H chemical shift (ppm)

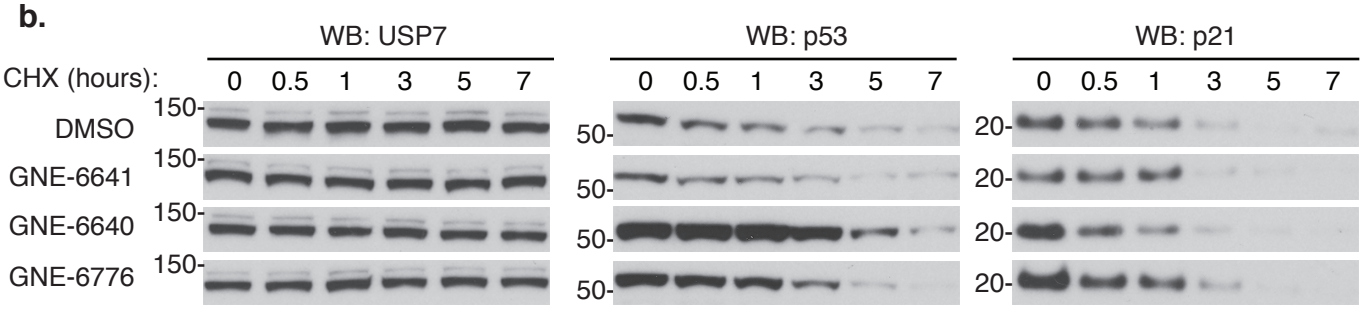

b.
reaction time (min):
TAMRA-peptide K48 Ub4 conjugate
TAMRA-peptide K63 Ub4 conjugate
0 5 10 15 20 25 30 35
peptide-K48 Ub4
peptide-K63 Ub4
peptide-K48 Ub3
peptide-K48 Ub2
free peptide
densitometry units (x10^6)
reaction time (min)
free peptide
peptide-Ub4
peptide-Ub3
peptide-Ub2

a.

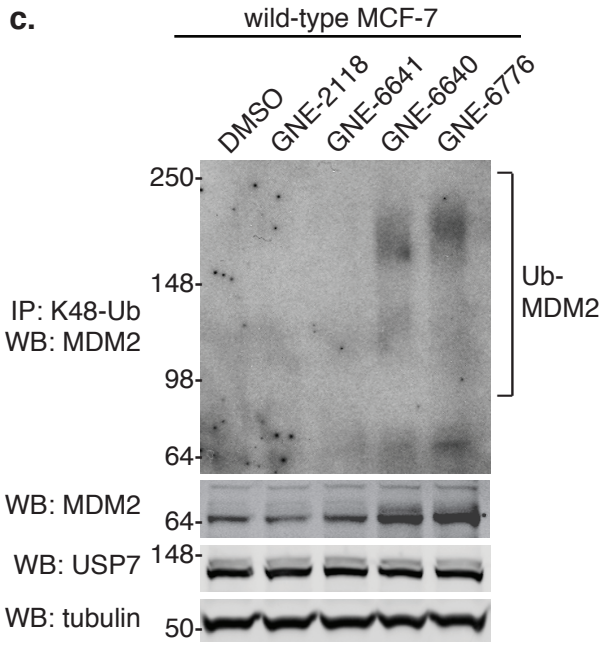
	active compounds			inactive control compounds	
Compound	GNE-2916	GNE-6640	GNE-6776	GNE-2118	GNE-6641
Full Length USP7 IC_{50} (μM)	2.63±0.43	0.75±0.37	1.34±0.24	>63.3	>63.3
USP7 Catalytic Domain IC_{50} (μM)	1.40±0.14	0.43±0.07	0.61±0.15	>200	>63.3
Full Length USP47 IC_{50} (μM)	>200	20.3±2.2	>200	>200	>200
Full Length USP5 IC_{50} (μM)	>200	>200	>200	>200	>200
Ub-MDM2 MSD (μM)	7.60±0.57	0.89±0.73	1.5	>50	>50



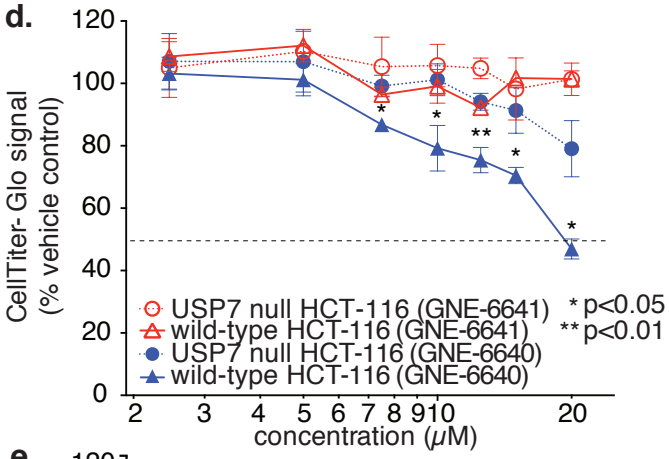
b.



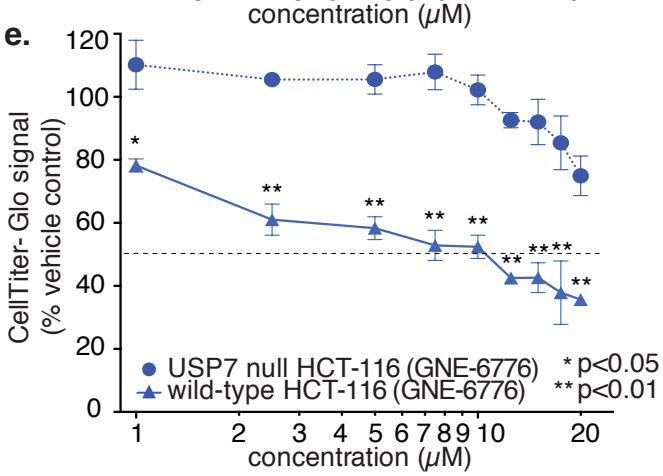
c.

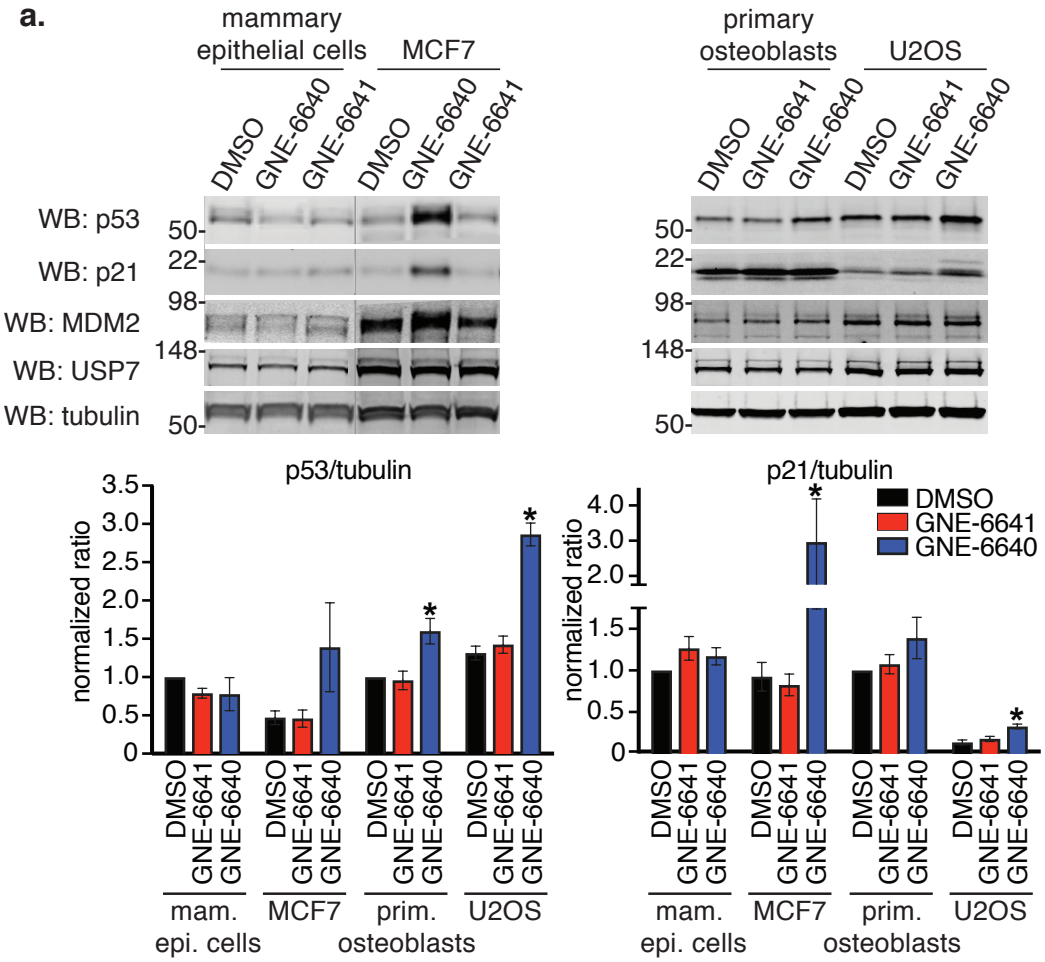


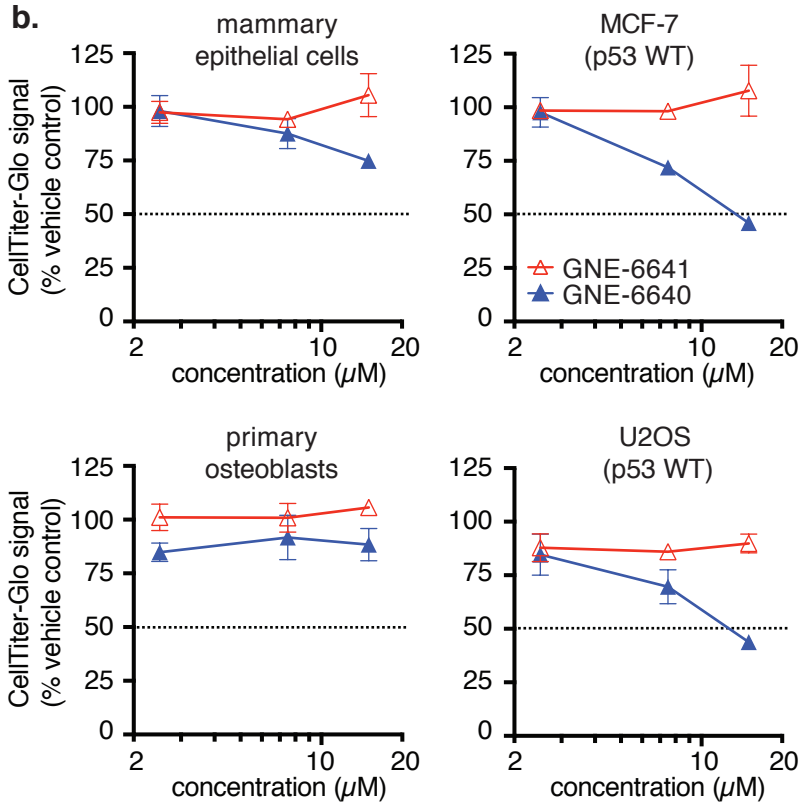
d.

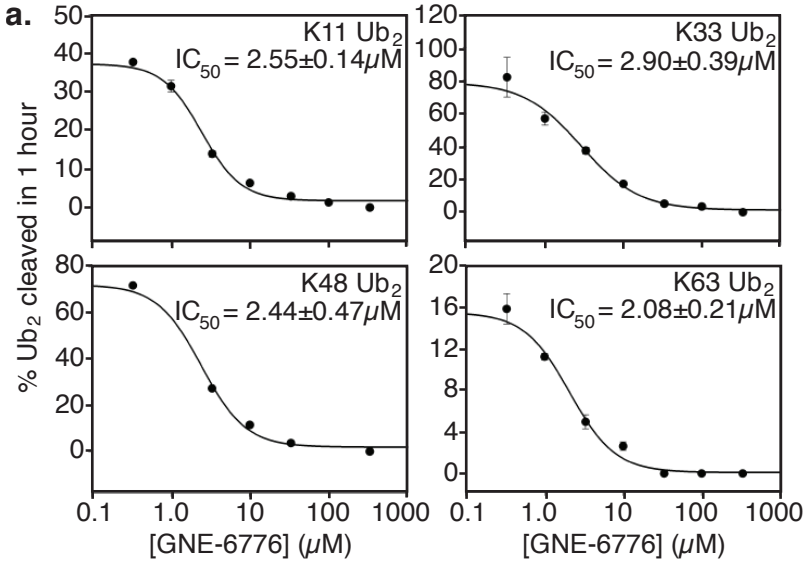


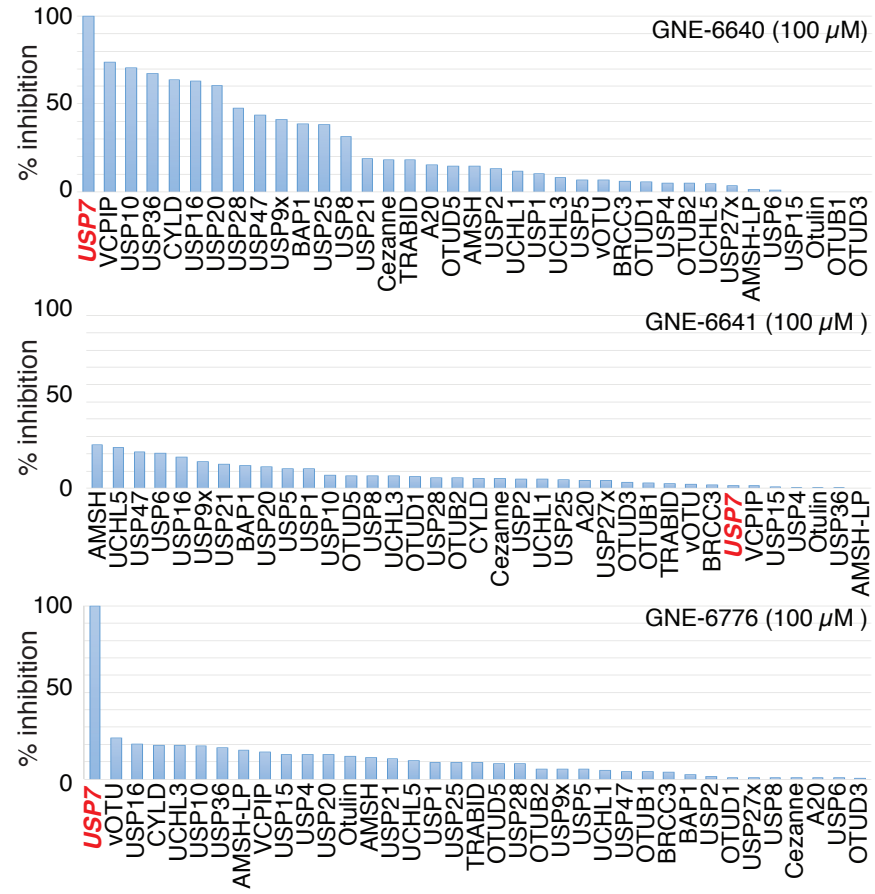
e.

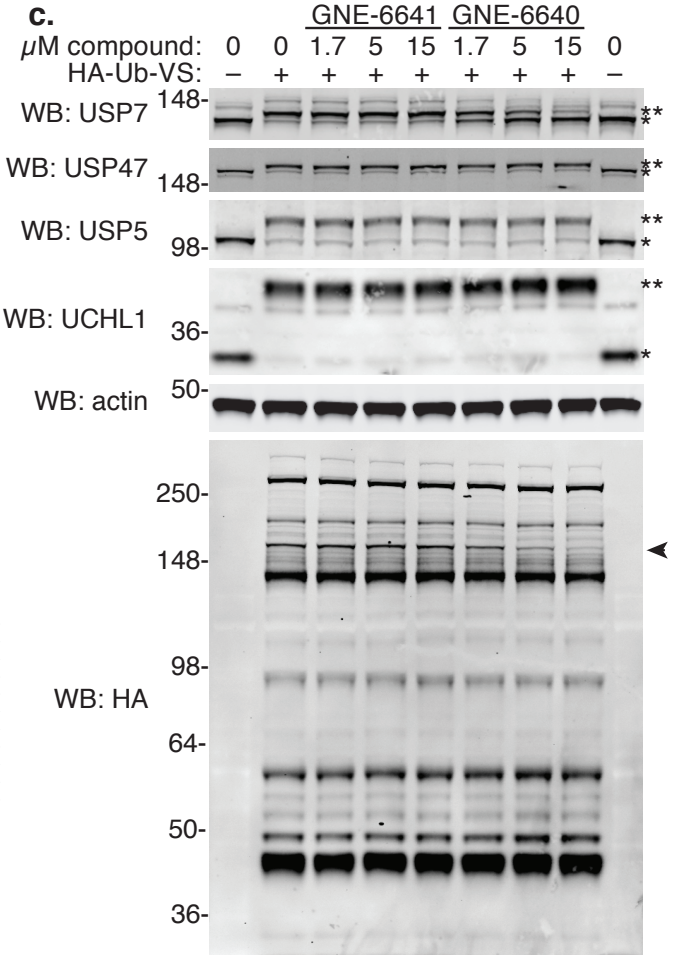


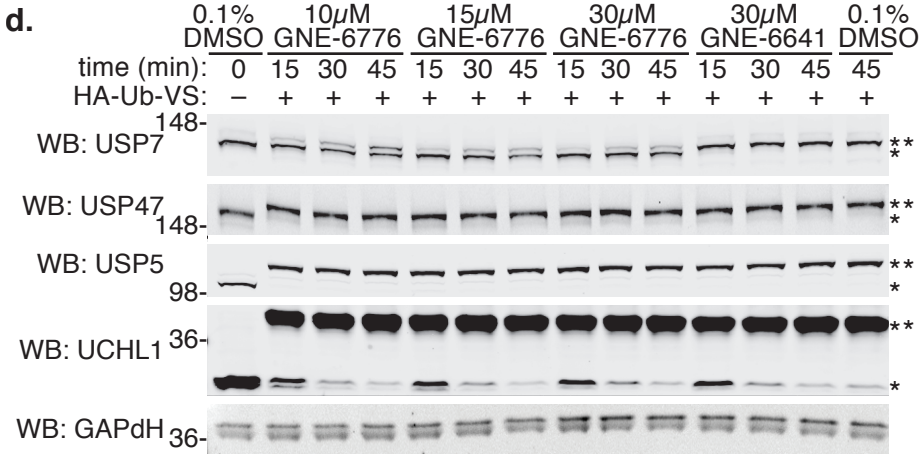

a.
mammary epithelial cells
MCF7
primary osteoblasts
U2OS
DMSO
GNE-6640
GNE-6641
WB: p53
WB: p21
WB: MDM2
WB: USP7
WB: tubulin
50-
22-
98-
148-
50-
p53/tubulin
p21/tubulin
normalized ratio
DMSO
GNE-6641
GNE-6640
mam. epi. cells
MCF7
prim. osteoblasts
U2OS

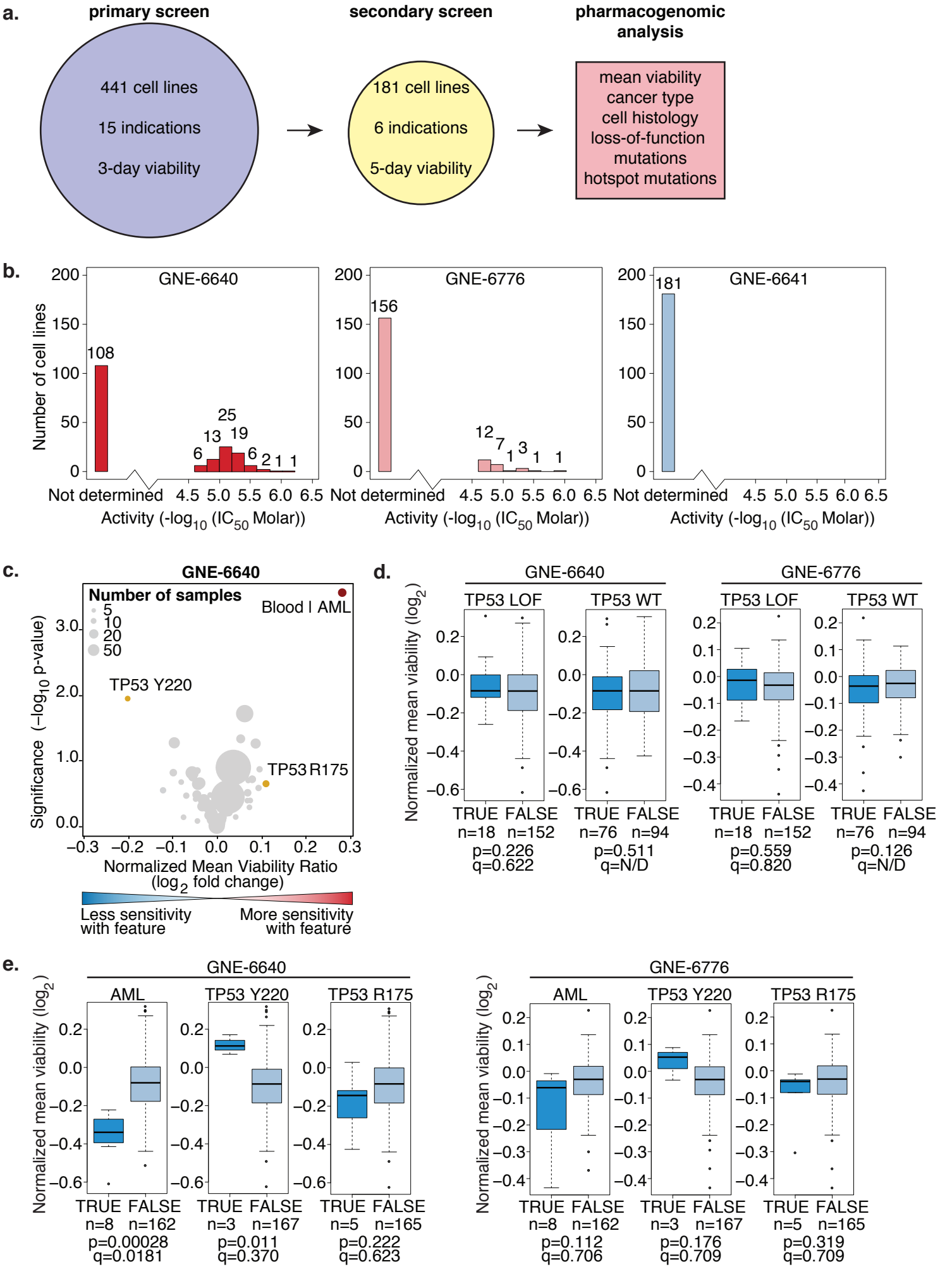

b.
mammary epithelial cells
MCF-7 (p53 WT)
primary osteoblasts
U2OS (p53 WT)
CellTiter-Glo signal (% vehicle control)
concentration (μM)
GNE-6641
GNE-6640

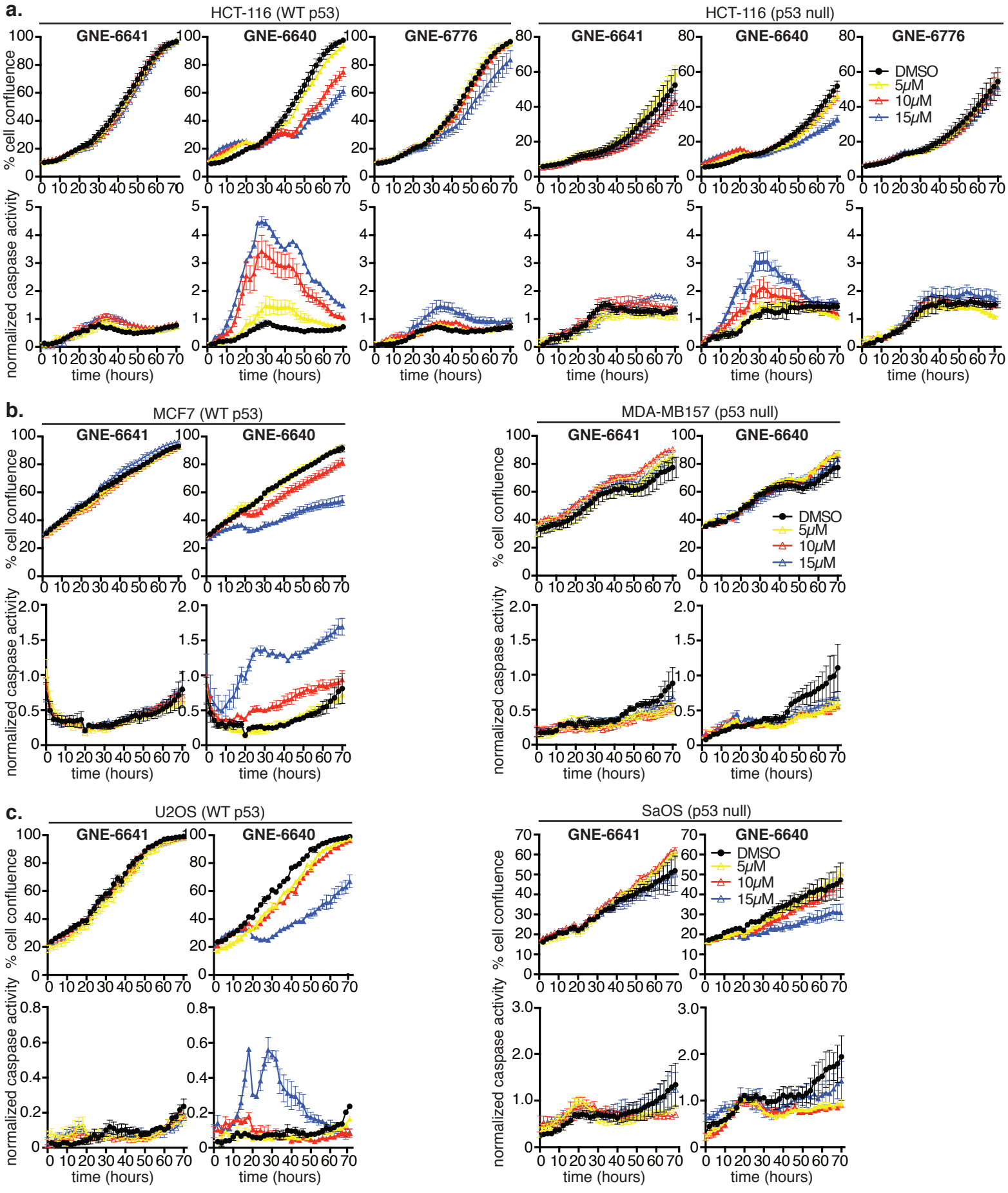

a.
K11 Ub2
IC50 = 2.55±0.14μM
K33 Ub2
IC50 = 2.90±0.39μM
K48 Ub2
IC50 = 2.44±0.47μM
K63 Ub2
IC50 = 2.08±0.21μM
% Ub2 cleaved in 1 hour
[GNE-6776] (μM)
[GNE-6776] (μM)

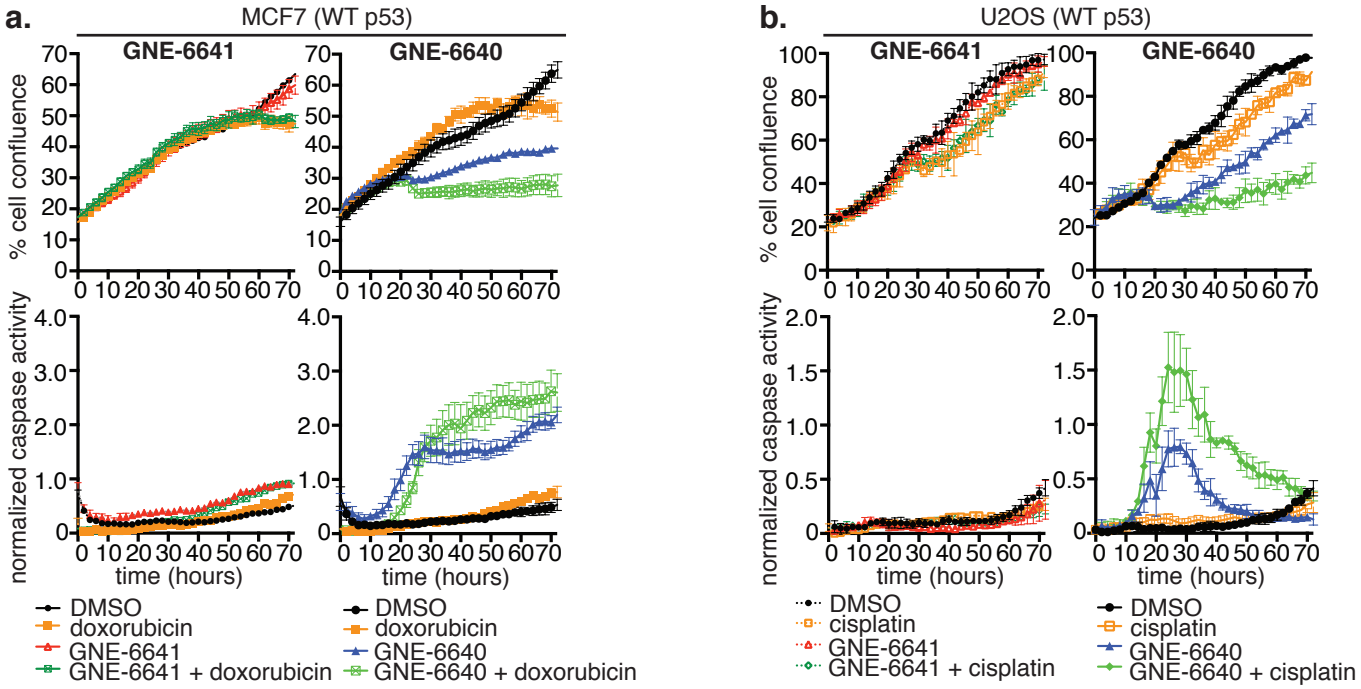

b.
GNE-6640 (100 μM)
GNE-6641 (100 μM)
GNE-6776 (100 μM)
% inhibition
USP7
VCPIP
USP10
USP36
CYLD
USP16
USP20
USP28
USP47
USP9x
BAP1
USP25
USP8
USP21
Cezanne
TRABID
A20
OTUD5
AMSH
USP2
UCHL1
USP1
UCHL3
USP5
vOTU
BRCC3
OTUD1
USP4
OTUB2
UCHL5
USP27x
AMSH-LP
USP6
USP15
Otulin
OTUB1
OTUD3
AMSH
UCHL5
USP47
USP6
USP16
USP9x
USP21
BAP1
USP20
USP5
USP1
USP10
OTUD5
USP8
UCHL3
OTUD1
USP28
OTUB2
CYLD
Cezanne
USP2
UCHL1
USP25
A20
USP27x
OTUD3
OTUB1
TRABID
vOTU
BRCC3
USP7
VCPIP
USP15
USP4
Otulin
USP36
AMSH-LP
USP7
vOTU
USP16
CYLD
UCHL3
USP10
USP36
AMSH-LP
VCPIP
USP15
USP4
USP20
Otulin
AMSH
USP21
UCHL5
USP1
USP25
TRABID
OTUD5
USP28
OTUB2
USP9x
USP5
UCHL1
USP47
OTUB1
BRCC3
BAP1
USP2
OTUD1
USP27x
USP8
Cezanne
A20
USP6
OTUD3

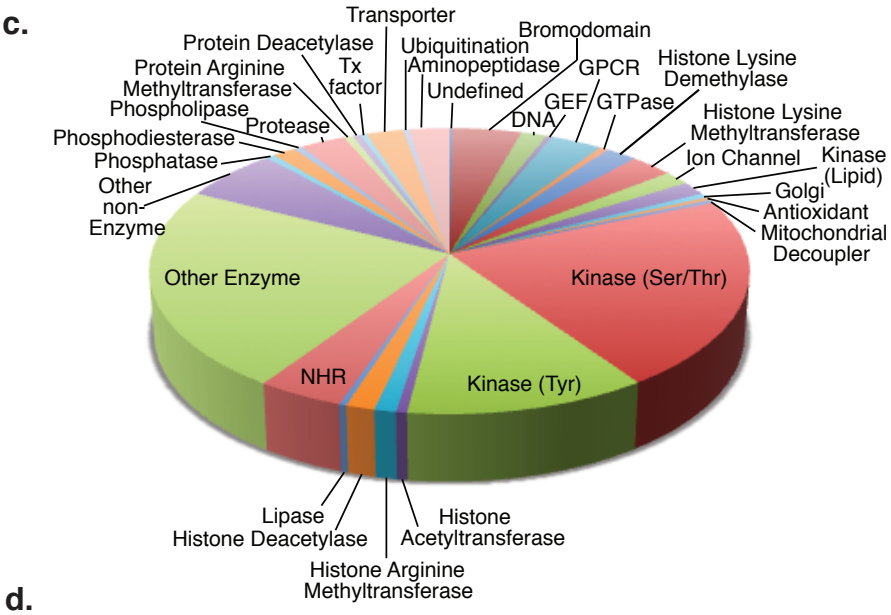

c.
GNE-6641
GNE-6640
μM compound: 0 0 1.7 5 15 1.7 5 15 0
HA-Ub-VS: – + + + + + + + –
WB: USP7
WB: USP47
WB: USP5
WB: UCHL1
WB: actin
WB: HA
148-
98-
36-
50-
64-
250-

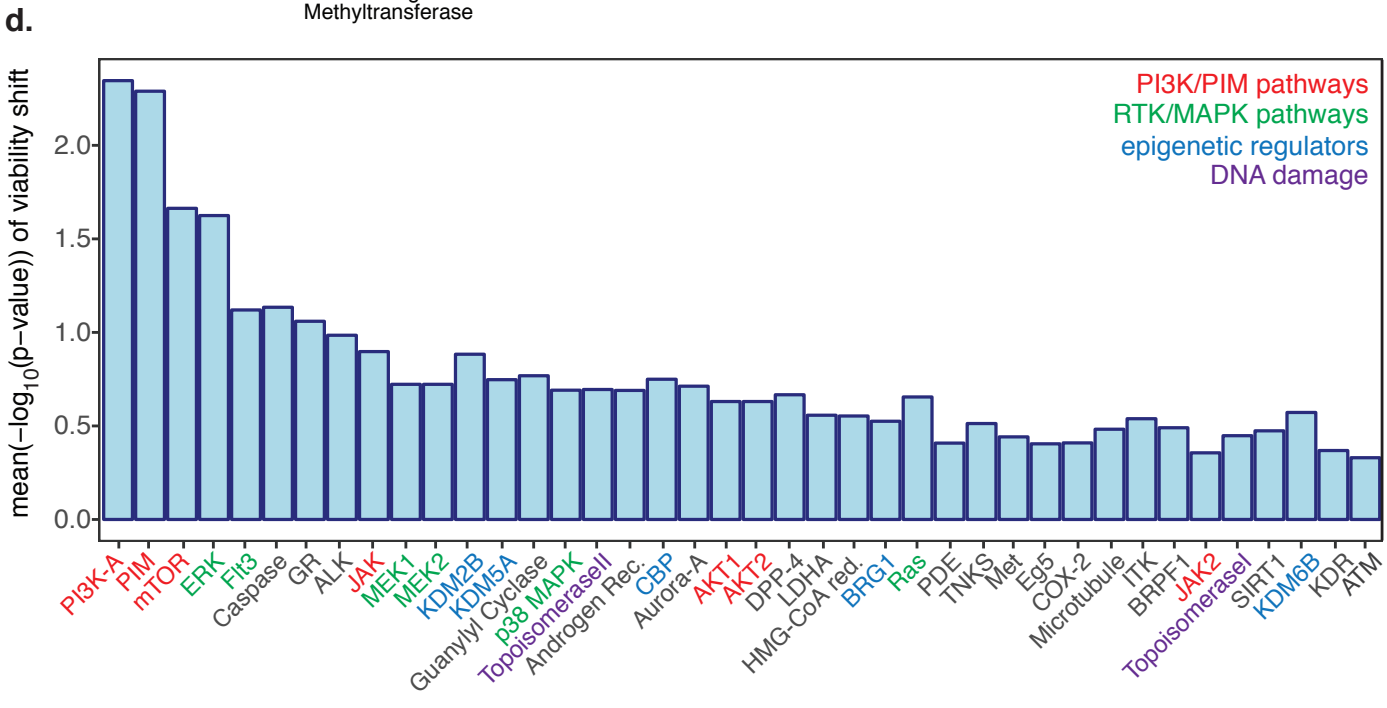

d.
0.1% DMSO
10μM GNE-6776
15μM GNE-6776
30μM GNE-6776
30μM GNE-6641
0.1% DMSO
time (min): 0 15 30 45 15 30 45 15 30 45 15 30 45 45
HA-Ub-VS: – + + + + + + + + + + + + +
WB: USP7
WB: USP47
WB: USP5
WB: UCHL1
WB: GAPdH
148-
98-
36-

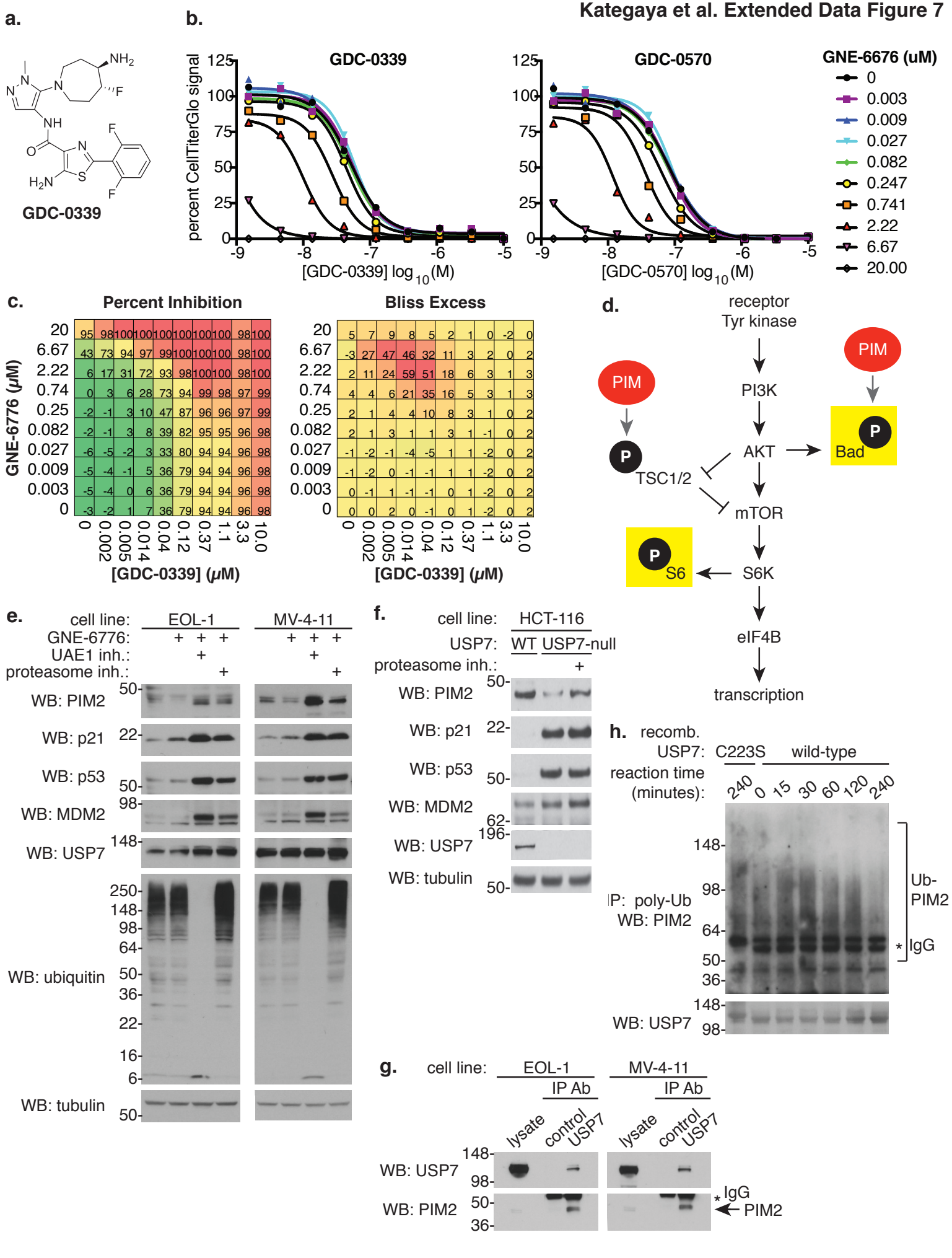

a.
primary screen
441 cell lines
15 indications
3-day viability
secondary screen
181 cell lines
6 indications
5-day viability
pharmacogenomic analysis
mean viability
cancer type
cell histology
loss-of-function mutations
hotspot mutations
b.
GNE-6640
GNE-6776
GNE-6641
Number of cell lines
Not determined
Activity (-log10 (IC50 Molar))
c.
GNE-6640
Number of samples
Blood | AML
TP53 Y220
TP53 R175
Significance (−log10 p-value)
Normalized Mean Viability Ratio (log2 fold change)
Less sensitivity with feature
More sensitivity with feature
d.
GNE-6640
GNE-6776
TP53 LOF
TP53 WT
Normalized mean viability (log2)
TRUE FALSE
n=18 n=152
p=0.226
q=0.622
n=76 n=94
p=0.511
q=N/D
p=0.559
q=0.820
p=0.126
q=N/D
e.
GNE-6640
GNE-6776
AML
TP53 Y220
TP53 R175
n=8 n=162
p=0.00028
q=0.0181
n=3 n=167
p=0.011
q=0.370
n=5 n=165
p=0.222
q=0.623
p=0.112
q=0.706
p=0.176
q=0.709
p=0.319
q=0.709


a.
HCT-116 (WT p53)
HCT-116 (p53 null)
GNE-6641
GNE-6640
GNE-6776
% cell confluence
normalized caspase activity
time (hours)
DMSO
5μM
10μM
15μM
b.
MCF7 (WT p53)
MDA-MB157 (p53 null)
c.
U2OS (WT p53)
SaOS (p53 null)


a.
MCF7 (WT p53)
GNE-6641
GNE-6640
% cell confluence
normalized caspase activity
time (hours)
DMSO
doxorubicin
GNE-6641
GNE-6641 + doxorubicin
DMSO
doxorubicin
GNE-6640
GNE-6640 + doxorubicin
b.
U2OS (WT p53)
GNE-6641
GNE-6640
% cell confluence
normalized caspase activity
time (hours)
DMSO
cisplatin
GNE-6641
GNE-6641 + cisplatin
DMSO
cisplatin
GNE-6640
GNE-6640 + cisplatin


c.
Transporter
Protein Deacetylase
Ubiquitination
Bromodomain
Protein Arginine Methyltransferase
Tx factor
Aminopeptidase
GPCR
Histone Lysine Demethylase
Undefined
Phospholipase
GEF
GTPase
DNA
Histone Lysine Methyltransferase
Protease
Phosphodiesterase
Ion Channel
Kinase (Lipid)
Phosphatase
Golgi
Other non-Enzyme
Antioxidant
Mitochondrial Decoupler
Other Enzyme
Kinase (Ser/Thr)
NHR
Kinase (Tyr)
Lipase
Histone Deacetylase
Histone Acetyltransferase
Histone Arginine Methyltransferase


d.
mean(−log10(p−value)) of viability shift
PI3K/PIM pathways
RTK/MAPK pathways
epigenetic regulators
DNA damage
PI3K-A
PIM
mTOR
ERK
Flt3
Caspase
GR
ALK
JAK
MEK1
MEK2
KDM2B
KDM5A
Guanylyl Cyclase
p38 MAPK
TopoisomeraseII
Androgen Rec.
CBP
Aurora-A
AKT1
AKT2
DPP-4
LDHA
HMG-CoA red.
BRG1
Ras
PDE
TNKS
Met
Eg5
COX-2
Microtubule
ITK
BRPF1
JAK2
TopoisomeraseI
SIRT1
KDM6B
KDR
ATM


a.
GDC-0339
b.
GDC-0339
GDC-0570
percent CellTiterGlo signal
[GDC-0339] log10(M)
[GDC-0570] log10(M)
GNE-6676 (uM)
0
0.003
0.009
0.027
0.082
0.247
0.741
2.22
6.67
20.00
c.
Percent Inhibition
Bliss Excess
GNE-6776 (μM)
[GDC-0339] (μM)
d.
receptor
Tyr kinase
PIM
PI3K
AKT
Bad
TSC1/2
mTOR
S6K
S6
eIF4B
transcription
e.
cell line:
EOL-1
MV-4-11
GNE-6776:
UAE1 inh.:
proteasome inh.:
WB: PIM2
WB: p21
WB: p53
WB: MDM2
WB: USP7
WB: ubiquitin
WB: tubulin
f.
cell line:
HCT-116
USP7:
WT
USP7-null
proteasome inh.:
WB: PIM2
WB: p21
WB: p53
WB: MDM2
WB: USP7
WB: tubulin
h.
recomb.
USP7:
C223S
wild-type
reaction time
(minutes):
IP: poly-Ub
WB: PIM2
Ub-
PIM2
IgG
WB: USP7
g.
cell line:
EOL-1
MV-4-11
IP Ab
lysate
control
USP7
WB: USP7
WB: PIM2
IgG
PIM2

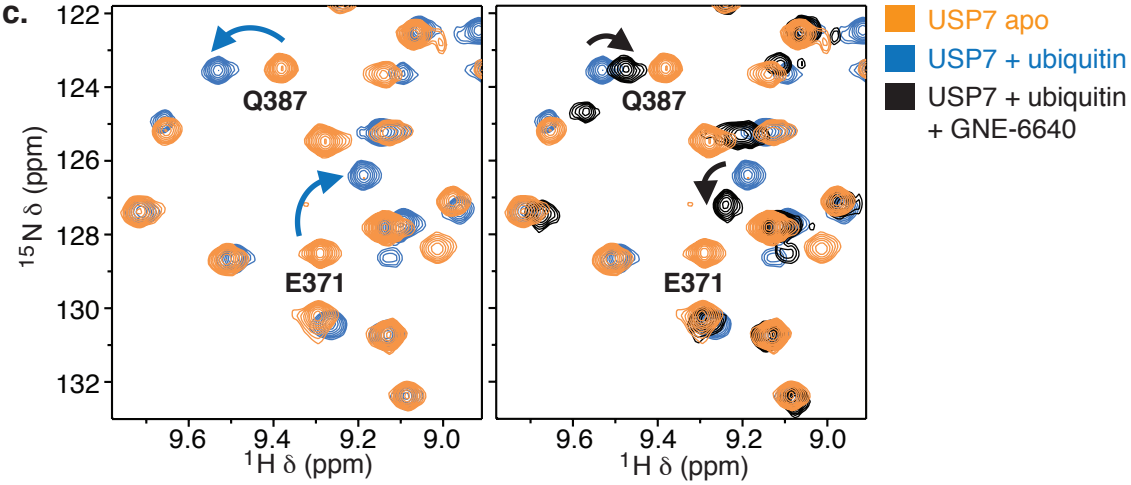
a.

USP7 inhibitor	K_M (μM)	K_{cat} (s^{-1})	K_{cat} / K_M (M^{-1} s^{-1})
GNE-6640	5.0 ± 0.5	0.22 ± 0.03	4.4 ± 0.7 x 10^4
GNE-6776	15.1 ± 0.7	0.061 ± 0.007	4.0 ± 0.5 x 10^3
GNE-6641	3.6 ± 0.1	1.36 ± 0.06	3.8 ± 0.2 x 10^5
no inhibitor	3.4 ± 0.1	1.54 ± 0.02	4.5 ± 0.1 x 10^5

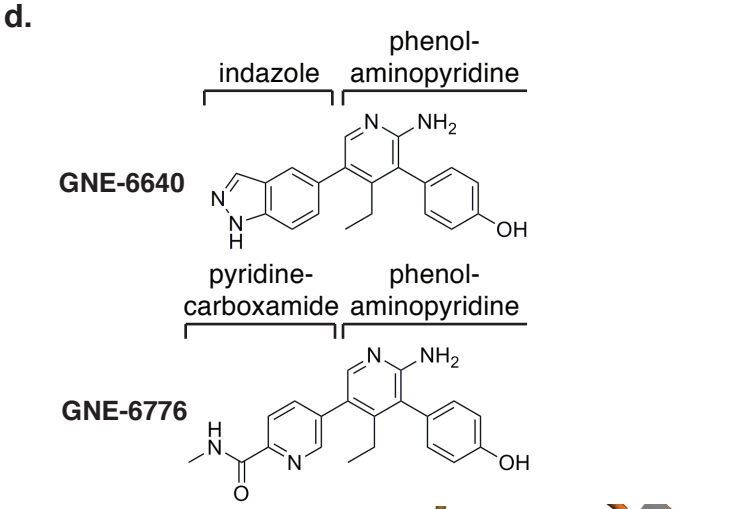
b.

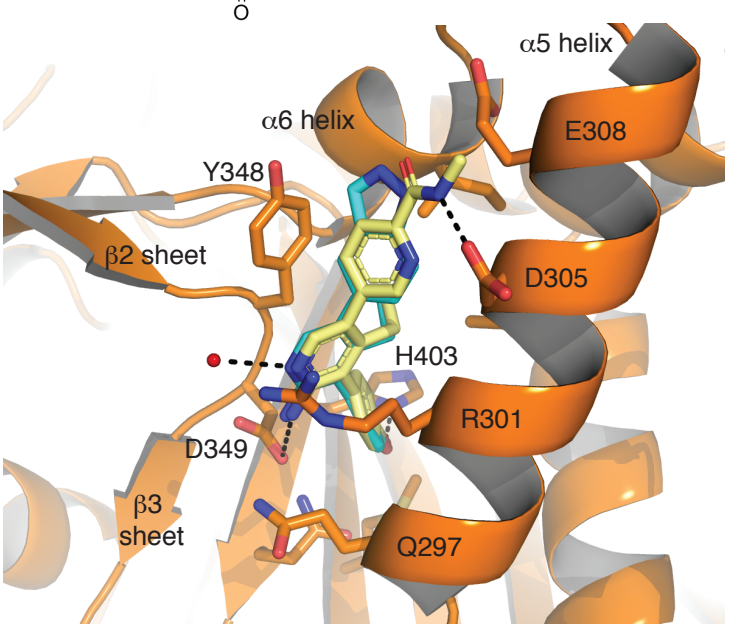
USP7 inhibitor	USP7/ubiquitin K_D (μM)
GNE-6640	372.7 ± 59.1
GNE-6776	246.0 ± 20.3
GNE-6641	156.3 ± 12.9
no inhibitor	133.4 ± 7.1

c.



d.





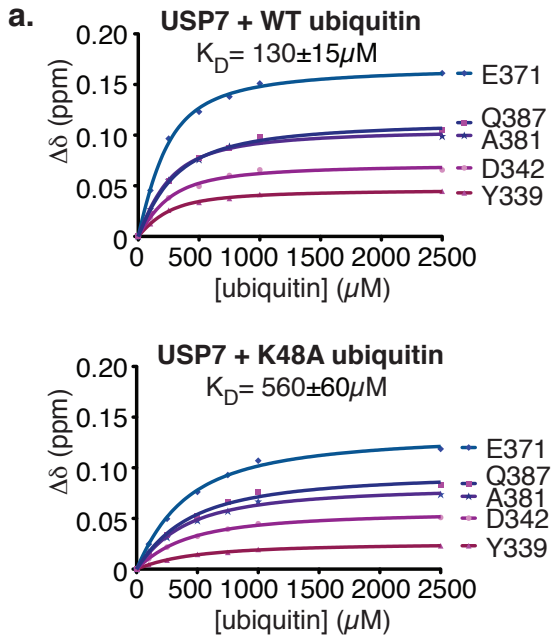
e.

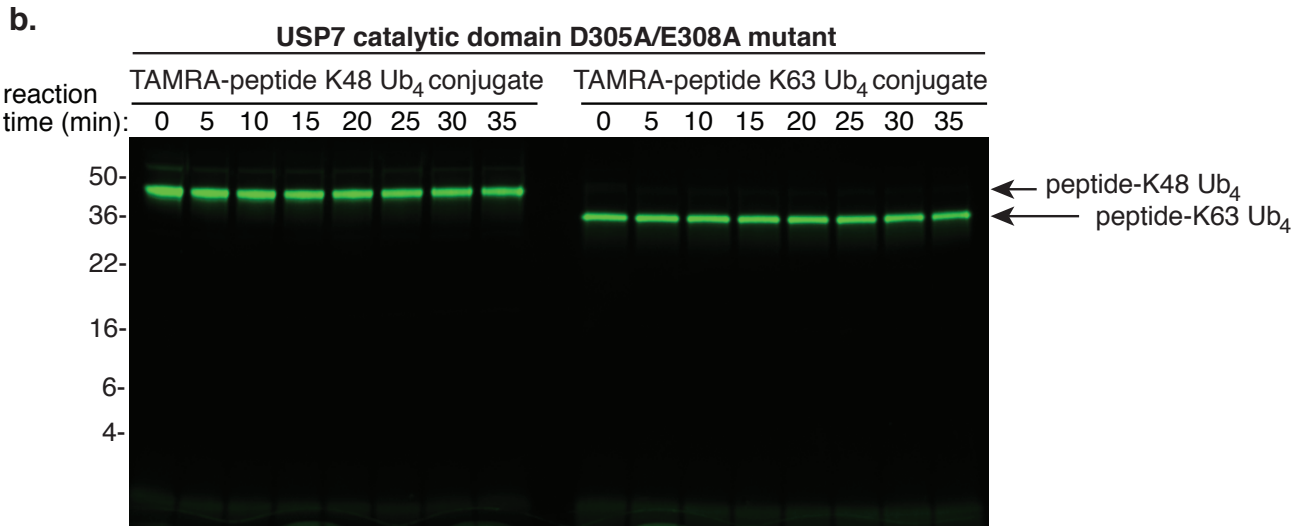
	GNE-6640 (PDB code 5UQV)	GNE-6776 (PDB code 5UQX)
Data collection		
Space group	P21	P21
Cell dimensions		
a, *b*, c (Å)	75.79 68.44 76.74	76.36 67.53 76.86
α, *β*, *γ* (°)	90 95.45 90	90 97 90
Resolution (Å)	76.4 -2.84 (2.85-2.84)[a]	67.8 - 2.23) (2.24 -2.23)
R_{merge}	0.052 (0.491)	0.18 (0.94)
$I/\sigma(I)$	15.7 (2.9)	9.1 (3.8)
$CC_{1/2}$	0.999 (0.87)	0.6 (0.016)
Completeness (%)	96.3 (98.5)	96.8 (96.0)
Redundancy	3.6 (3.8)	3.8 (4.0)
Refinement		
Resolution (Å)	43.5-2.84	38.2-2.23
No. reflections	17844	37714
R_{work}/R_{free}	0.221/0.278	0.19/0.223
No. atoms		
Protein	5384	5535
Ligand/ion	52	52
Water	2	221
B factors		
Protein	74	54
Ligand/ion	66	41
Water		48
R.m.s. deviations		
Bond lengths (Å)	0.001	0.009
Bond angles (°)	0.5	1.0

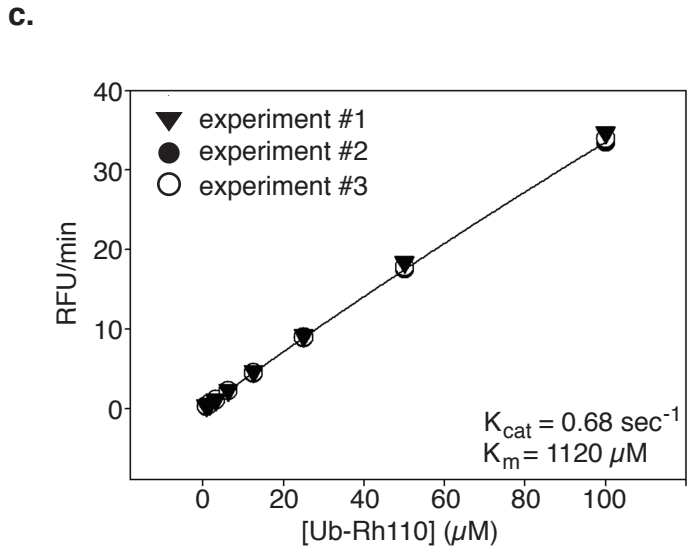
[a]Values in parentheses are for highest resolution shell

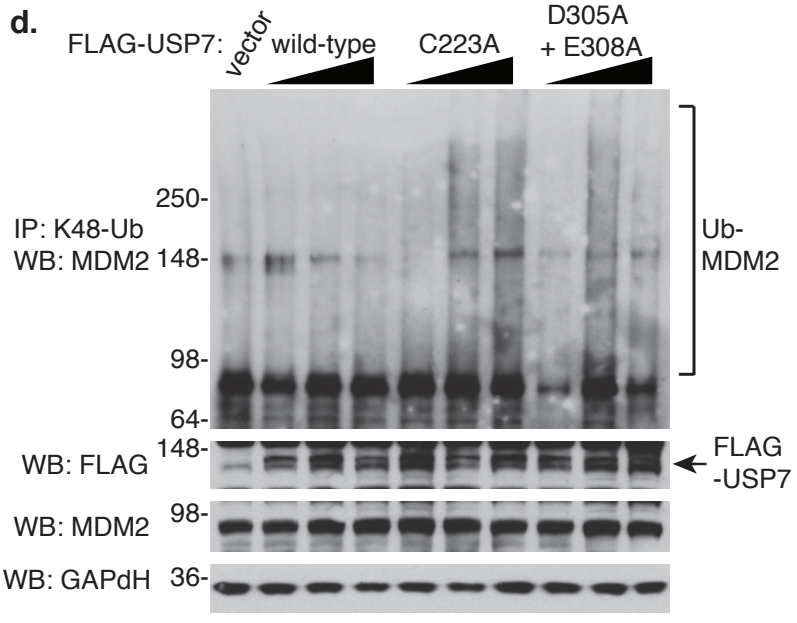
Molprobity score for GNE-6640 = 2.53

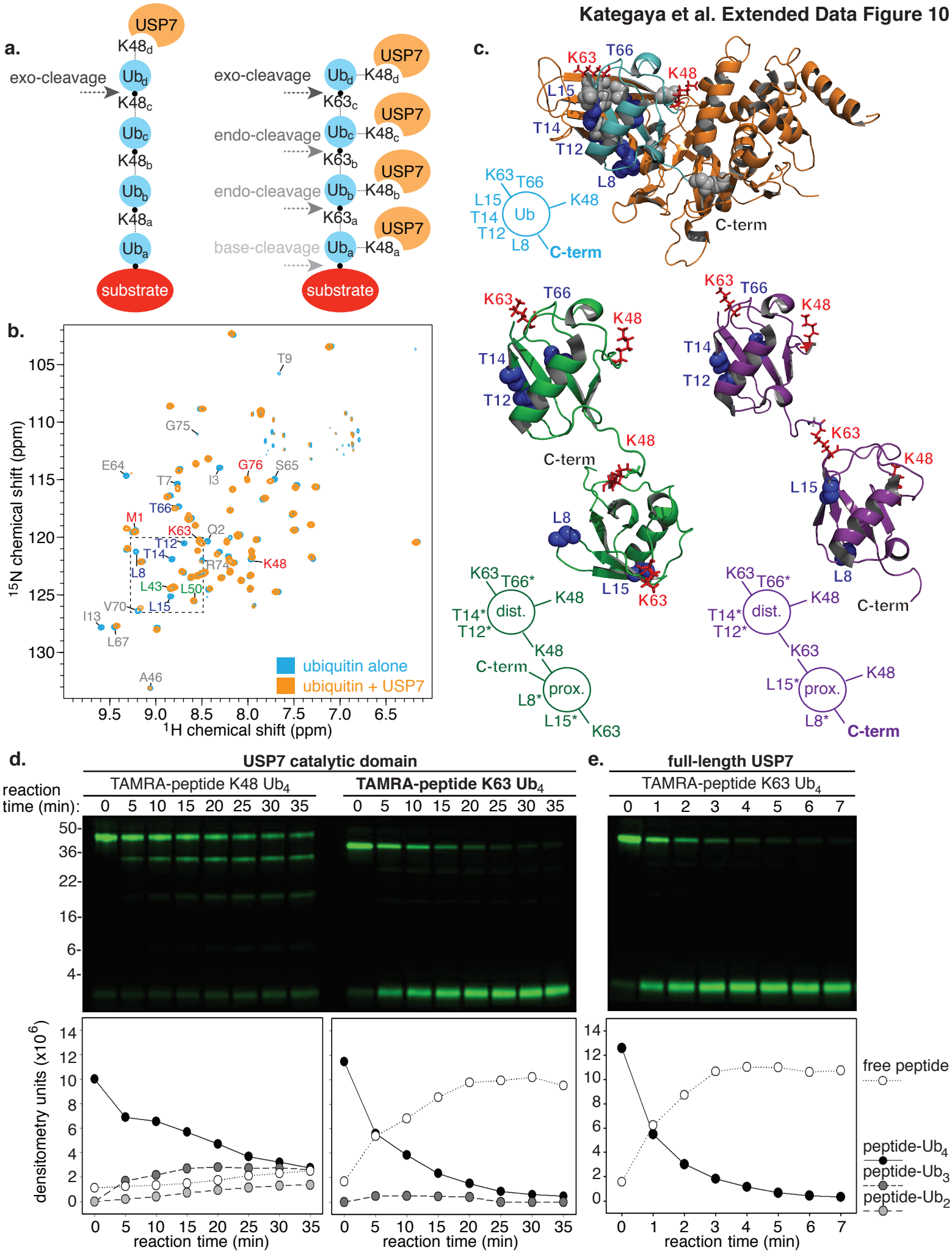
Molprobity score for GNE-6776 = 2.42


a.
USP7 + WT ubiquitin
K_D= 130±15μM
E371
Q387
A381
D342
Y339
Δδ (ppm)
[ubiquitin] (μM)
USP7 + K48A ubiquitin
K_D= 560±60μM
E371
Q387
A381
D342
Y339
Δδ (ppm)
[ubiquitin] (μM)

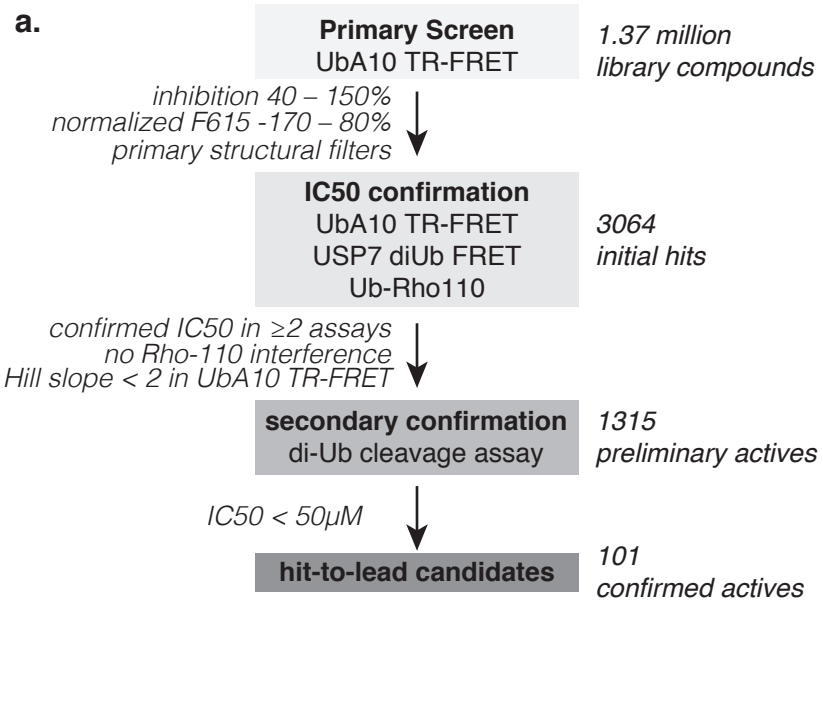

b.
USP7 catalytic domain D305A/E308A mutant
TAMRA-peptide K48 Ub4 conjugate
TAMRA-peptide K63 Ub4 conjugate
reaction time (min):
0 5 10 15 20 25 30 35
0 5 10 15 20 25 30 35
50-
36-
22-
16-
6-
4-
peptide-K48 Ub4
peptide-K63 Ub4


c.
experiment #1
experiment #2
experiment #3
RFU/min
K_cat = 0.68 sec-1
K_m= 1120 μM
[Ub-Rh110] (μM)

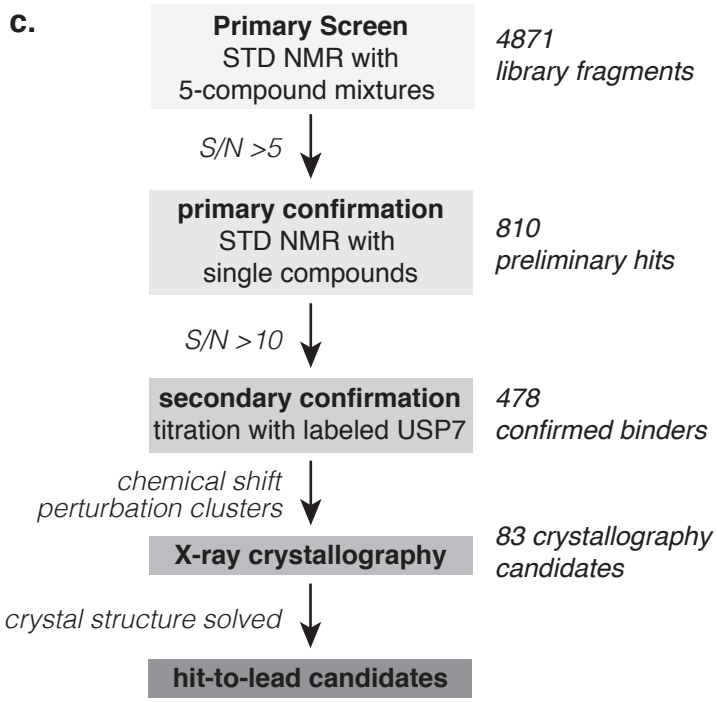

d.
FLAG-USP7:
vector
wild-type
C223A
D305A + E308A
IP: K48-Ub
WB: MDM2
250-
148-
98-
64-
Ub-MDM2
WB: FLAG
148-
FLAG-USP7
WB: MDM2
98-
WB: GAPdH
36-


a.
USP7
exo-cleavage
endo-cleavage
base-cleavage
substrate
b.
ubiquitin alone
ubiquitin + USP7
15N chemical shift (ppm)
1H chemical shift (ppm)
c.
C-term
d.
USP7 catalytic domain
TAMRA-peptide K48 Ub4
TAMRA-peptide K63 Ub4
reaction time (min):
densitometry units (x10^6)
reaction time (min)
e.
full-length USP7
free peptide
peptide-Ub4
peptide-Ub3
peptide-Ub2

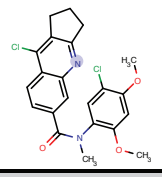
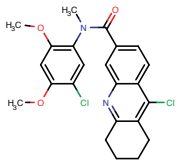
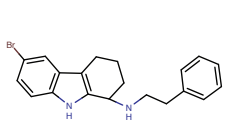
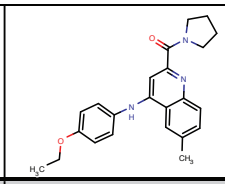
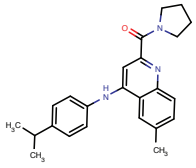
a.



c.



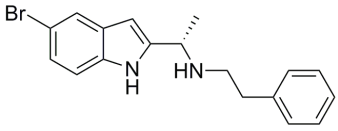
b.

					
Cluster	1	1	2	3	3
% Inhibition Primary Screen	76.8	76.3	66.4	53.3	50.1
UbA10 IC_{50} (μM)	1.17	1.37	> 49.5	2.58	2.79
diUb IC_{50} (μM)	0.36	0.49	6.42	2.49	1.53
Ub-Rho110 IC_{50} (μM)	0.29	0.31	2.16	2.65	N/D
MS Activity IC_{50} (μM)	0.13	0.15	14.87	0.5	0.61
MS Activity Hill Slope	0.83	1.06	1.02	0.87	1.12

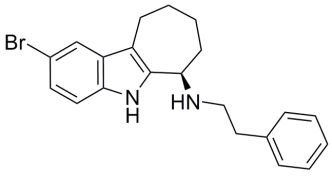
N/D = not determined

Kategaya et al. Supplementary Information Figure 2

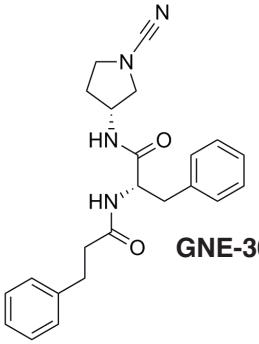
	Indole tricyclic		Peptidomimetic	Tetrahydroacridine		Fragment	
Compound	**GNE-0300**	**GNE-8735**	**GNE-3086**	**GNE-2090**	**GNE-6831**	**GNE-2916**	
Full Length USP7 IC_{50} (μM)	1.03	0.47±0.39	8.80±5.20	4.64±1.38	0.75±1.30	2.63±0.43	**USP7 inhibition and selectivity**
USP7 Catalytic Domain IC_{50} (μM)	1.40	0.50±0.31	13.80	41.7±30.6	>200	1.40±0.18	
Full Length USP5 IC_{50} (μM)	160	>200	>100	>200	N/D	>200	
Full Length USP47 IC_{50} (μM)	140	>200	7.0	63.3	>200	>200	
Total MDM2 increase (IF)	none >20μM	increase 2.9μM	none >20μM	none >20μM	none >20μM	none >20μM	**cellular MDM2 assays**
Ub-MDM2 MSD (μM)	not done	>50.00	9.65±3.18	3.55±1.91	3.20±2.69	7.60±0.57	
Cathepsin-B Inhibition (μM)	> 40.00	6.12	not done	not done	> 40.00	not done	**protease selectivity**
Caspase-3 Inhibition	apparent precipitator	apparent precipitator	not done	not done	not done	not done	
Dynamic Light Scattering	pass	pass	pass	pass	pass	not done	**biophysical analysis**
Covalent Modification	not done	not done	yes	not done	yes	not done	



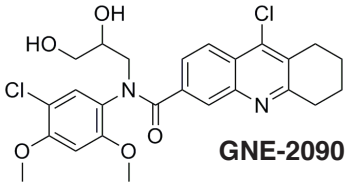
GNE-0300



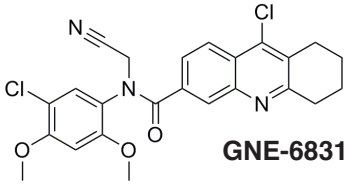
GNE-8735



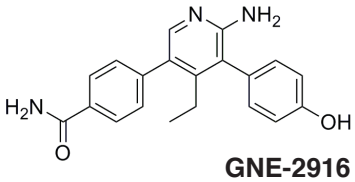
GNE-3086



GNE-2090

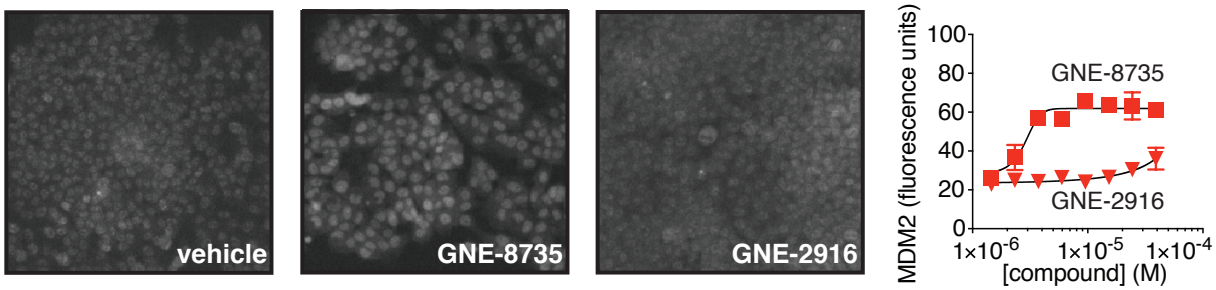


GNE-6831

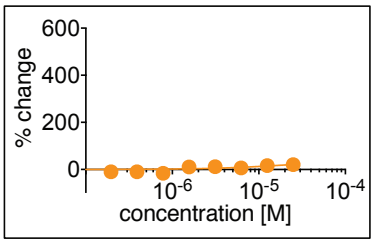
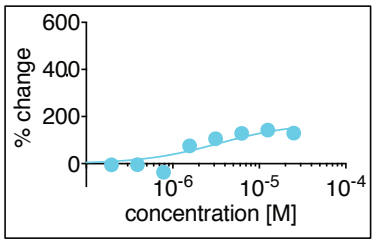
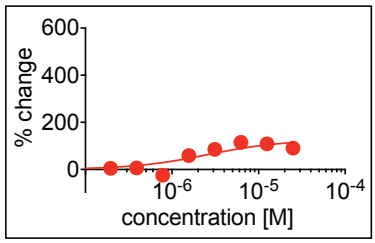
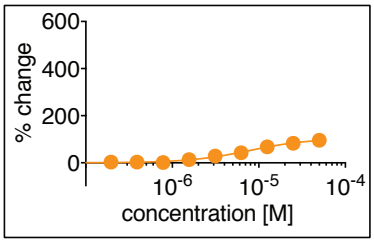
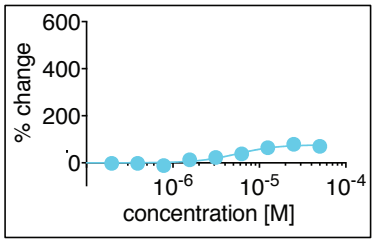
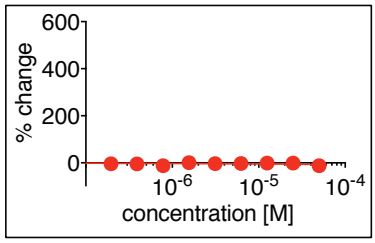
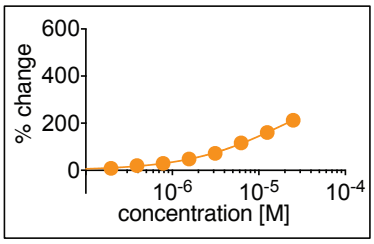
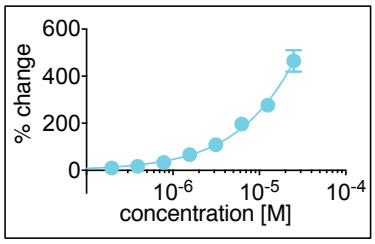
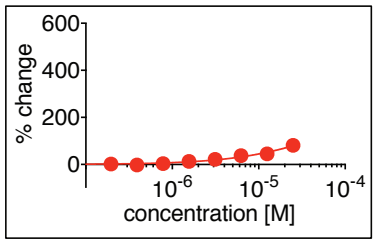
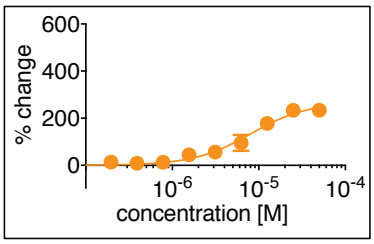
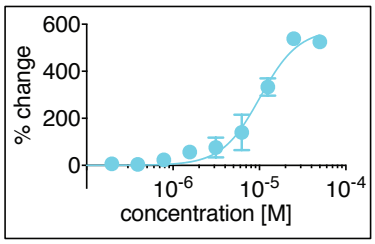
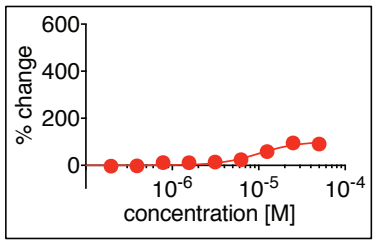
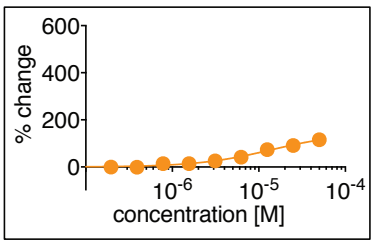
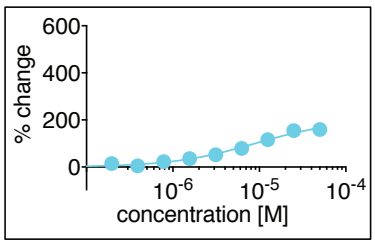
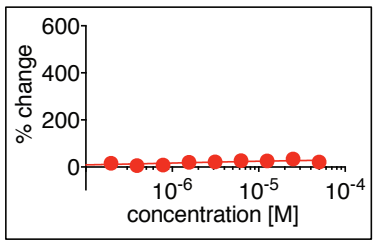


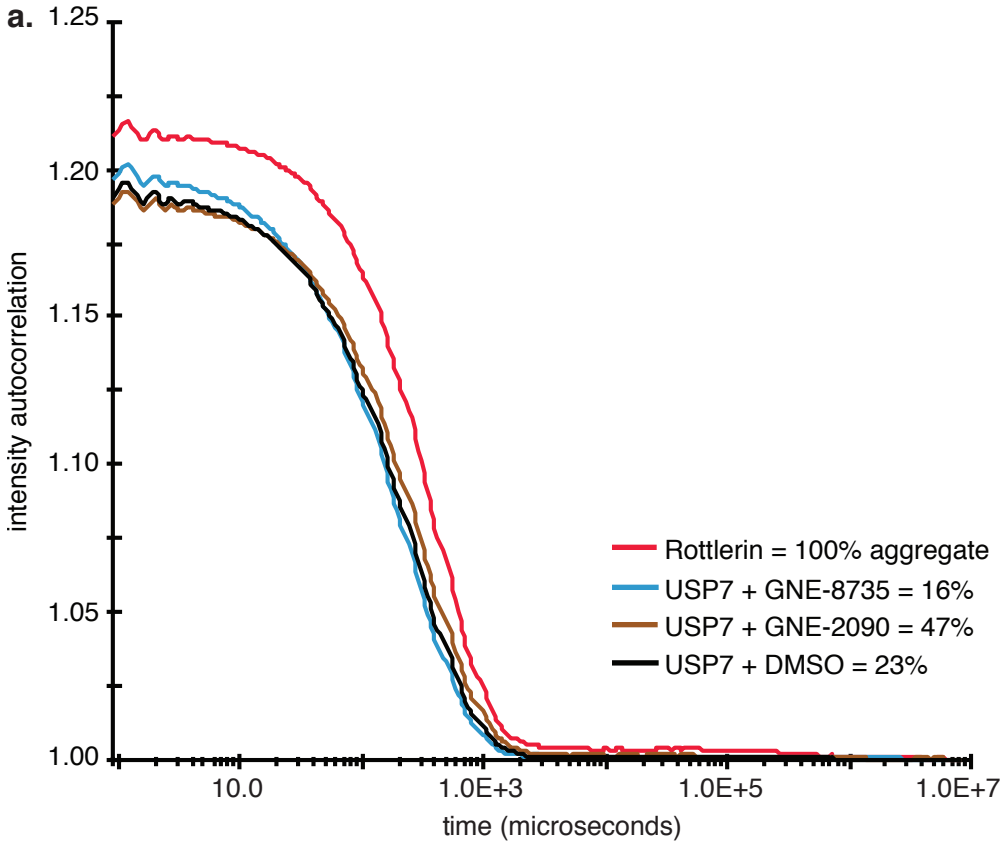
GNE-2916

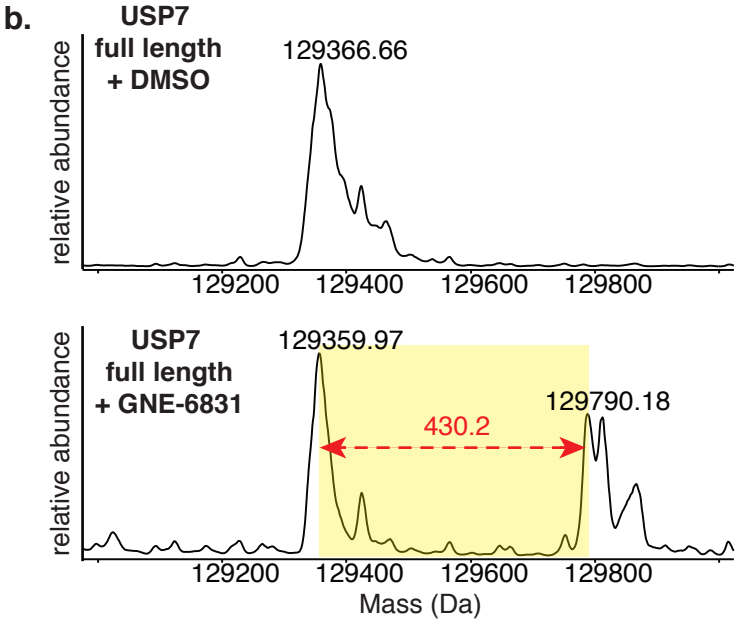
a.



b.

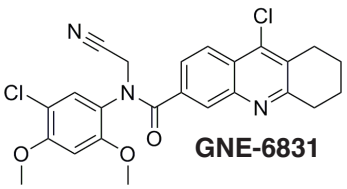
compound and EC_{50}	ratio Ub-MDM2/total MDM2	Ub-MDM2	total MDM2
GNE-8735 EC_{50} **>50μM**			
GNE-3086 EC_{50} **= 9.65±3.18μM**			
GNE-2090 EC_{50} **= 3.55±1.91μM**			
GNE-6831 EC_{50} **= 3.2±2.69μM**			
GNE-2916 EC_{50} **= 7.6±0.57μM**			

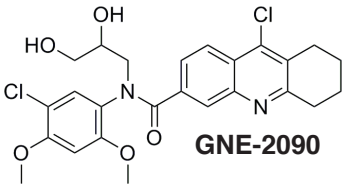

a.
1.25
1.20
1.15
1.10
1.05
1.00
intensity autocorrelation
10.0
1.0E+3
1.0E+5
1.0E+7
time (microseconds)
Rottlerin = 100% aggregate
USP7 + GNE-8735 = 16%
USP7 + GNE-2090 = 47%
USP7 + DMSO = 23%

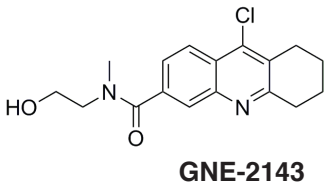

b.
USP7
full length
+ DMSO
129366.66
relative abundance
129200
129400
129600
129800
USP7
full length
+ GNE-6831
129359.97
129790.18
430.2
relative abundance
129200
129400
129600
129800
Mass (Da)

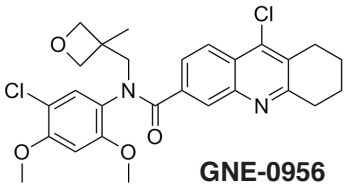
a.

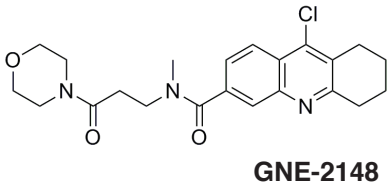
	active compounds		inactive control compounds		
Compound	**GNE-6831**	**GNE-2090**	**GNE-2148**	**GNE-2143**	**GNE-0956**
Full Length USP7 IC_{50} (μM)	0.75±0.13	4.64±1.38	>200	196	>20
USP7 Catalytic Domain IC_{50} (μM)	>100	41.7±30.6	>200	>200	>63
Full Length USP47 IC_{50} (μM)	>200	>200	>200	N/D	>200
Full Length USP5 IC_{50} (μM)	N/D	>200	>20	>20	>63
Ub-MDM2 MSD (μM)	30.8±11.1*	24.2±12.3*	N/D	N/D	N/D





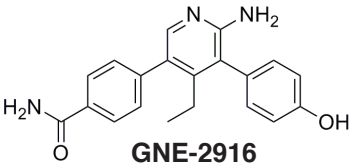


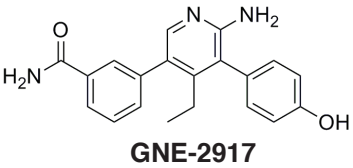


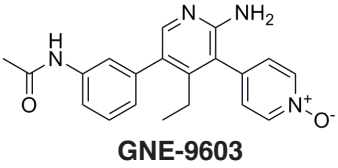


b.

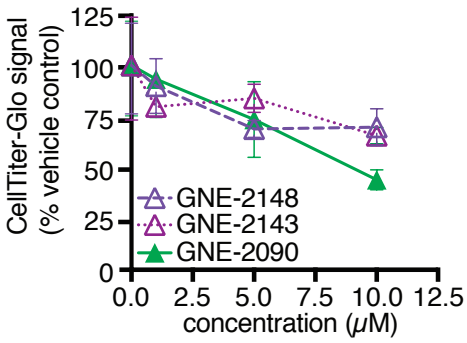
	active compound	inactive control compounds		
Compound	**GNE-2916**	**GNE-2931**	**GNE-9603**	**GNE-2917**
Full Length USP7 IC_{50} (μM)	2.63±0.43	>200	>200	114
USP7 Catalytic Domain IC_{50} (μM)	1.40±0.14	N/D	>200	66.8
Full Length USP47 IC_{50} (μM)	>200	>63.3	109	>200
Full Length USP5 IC_{50} (μM)	>200	N/D	>200	>20
Ub-MDM2 MSD (μM)	7.60±0.57	N/D	N/D	N/D

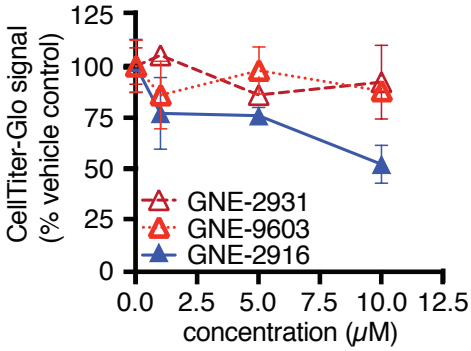




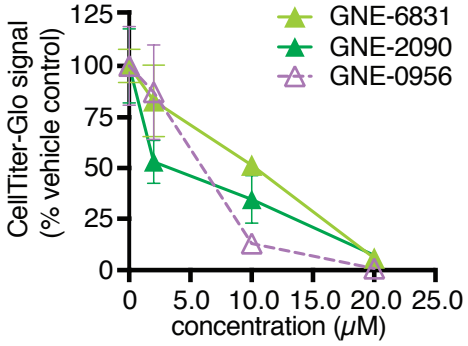


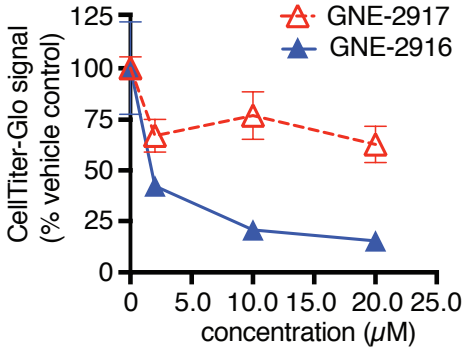
c. **AMO-1 cells**





d. **KMS12-PE cells**





a.

	GNE-6640	GNE-6776
MW/LogD$_{7.4}$/tPSA	330.38/4.40/87.0	348.40/3.30/3.4/101
LM CLhep (h/r/m/d/c)	14/46/75/24/33	4.1/15/37/14/12
Hep CLhep (h/r/m/d/c)	16/45/68/24/39	9.4/39/44/18/33
PPB % (h/r/m)	99.6/98.7/99.0	97.3/93.4/93.4
MDCK P_{app} ratio (B to A, A to B)	0.83 (12.17, 14.7)	0.75 (12.06, 16.10)